Mapping the social class structure

From occupational mobility to social class categories using network analysis

Toubøl, Jonas; Larsen, Anton Grau

Published in:
Sociology

DOI:
10.1177/0038038517704819

Publication date:
2017

Document version
Peer reviewed version

Mapping the social class structure: From occupational mobility to social class categories using network analysis

Abstract

This article develops a new explorative method for deriving social class categories from patterns of occupational mobility. In line with Max Weber, our research is based on the notion that, if class boundaries do not inhibit social mobility then the class categories are of little value. Thus, unlike dominant, theoretically defined class schemes, this paper derives social class categories from observed patterns in a mobility network covering intra-generational mobility. The network is based on a mobility table of 109 occupational categories tied together by 1,590,834 job shifts on the Danish labour market 2001-07. The number of categories are reduced from 109 to 34 by applying a new clustering algorithm specifically designed for the study of mobility tables (MONECA). These intra-generational social class categories are related to the central discussions of gender, income, education, and political action by providing empirical evidence of strong patterns of intra-generational class divisions along these lines.

Keywords

Class analysis, cluster analysis, intra-generational social mobility, Methodology, MONECA, social network analysis, Social class, Class theory, occupational mobility

Authors

Jonas Toubøl, University of Copenhagen, Department of Sociology

Anton Grau Larsen, Copenhagen Business School, Department of Business and Politics

Corresponding author

Jonas Toubøl, University of Copenhagen, Department of Sociology, Øster Farimagsgade 5, Bld. 16, DK-1014 Copenhagen K, Denmark, jt@soc.ku.dk

1. Introduction

This article proposes a new methodology for identifying class boundaries on the basis of a network of mobility between occupations. By deriving the class structure from mobility patterns we are able to stay closer to the theoretical ambition in Weberian class analysis, in which mobility between classes should be rare. This new method therefore contributes to the question of mapping the class structure which have preoccupied sociologists for more than a century (Marx and Engels, 2008 [1848]; Weber, 1978 [1922]). The questions of class boundaries and the correct empirical delimitation of classes are central to this debate. The problem of class boundaries is posed either as a question of the empirical operational outcome of competing class theories (Goldthorpe, 2007; Wright, 2005), as a question of the categories designed to tests various hypotheses about the structure of social mobility (e.g. Blau and Duncan, 1967; Clogg and Goodman, 1984) or as a question of the relative delimitation of clusters of individuals sharing tastes and dispositions (Bourdieu, 1987). Lately, this debate has changed profoundly, as the relevance of class theory, and indeed the very existence of classes has been brought into question (Pakulski, 2005).

Grusky, Weeden, and associates have set out to reconstruct class analysis by focusing on the realistic occupational categories expressing the division of labour as the site of production of class based behaviour and mobility patterns (Weeden and Grusky, 2005, 2012). What is at stake in this long running debate (Erikson et al., 2012; Grusky and Weeden, 2002) is whether class is best explained on the basis of 88 micro-classes (Jonsson et al., 2009), 11 meso-classes (Erikson and Goldthorpe, 1993) or 12 relational classes (Wright, 1989). Such theoretically derived class schemes are the most prominent, but they have been challenged by more data sensitive and descriptive approaches (Savage et al., 2013, 2015).

In the descriptive tradition, this study will demonstrate the viability of a data sensitive approach that derives intra-generational class structure from the mobility patterns between occupations. This is achieved by applying a new clustering algorithm, the *Mobility Network Clustering Algorithm* (MONECA), which we have developed[1] to register-data drawn from official records covering the entire Danish working population in the period 2001-2007. The result is 34 social classes reflecting the intra-generational social mobility structure. This result can be characterized as a mix of meso- and micro-classes, suggesting that the Danish intra-generational social class structure is not captured by any of the dominant universal class schemes. Rather, analyses of the distribution of income, education, political action and gender indicates that no single factor alone can account for the intra-generational social class structure revealed. This suggests that future research should investigate how different factors in different ways are involved in the formation of the various social classes of the intra-generational mobility structure.

Irrespective of how different theories have derived class boundaries, a common assumption is that these barriers also represent barriers to social mobility. Indeed, social mobility is at the very heart of class analysis. The power of class as an analytical concept consists in its ability to reflect that members of society cannot easily leave their class position. This entails that those sharing a class position will tend to perceive, and act upon the world in similar ways because their lives are conditioned by the same inescapable structures (Giddens, 1973). In truth, unless they are accompanied by barriers to mobility, class boundaries are only academic abstractions of little practical importance. The profound relation between social mobility and class finds an early and clear formulation in Max Weber's famous descriptive definition of social class: 'A "*social* class" makes up the totality of those class situations within which individual and generational mobility is easy and typical.' (Weber, 1978: 302 [1922]). While both Giddens and Weber place barriers to mobility at the core of their class theory, neither of them had methods that allowed mobility to define class boundaries. This article hopes to provide such a methodology.

None of today's dominant class schemes take their starting point in the observation of social mobility. The class schemes mentioned above all start from theoretical considerations about class formation, while mobility only plays a role as a, however privileged, variable to test the class schemes. Bourdieu starts from a theory of distinction, but his approach is more empirically sensitive and identifies classes by empirical analysis of latent dimensions or fields (Bourdieu, 1984). However, no observation of patterns of mobility are included in determining the class categories (Andersen and Hansen, 2012; Wacquant, 2013). The same is the case with the CAMSIS-scale, based on observed social interaction patterns of friendship and marriage (Prandy and Lambert, 2003; Stewart et al., 1980). If we acknowledge the fundamental weakness of any class scheme which does not reflect actual barriers to social mobility, the contemporary lack of such schemes challenges us to develop a method by which we can derive a class scheme from the observation of social mobility patterns between occupations. Such a description of the mobility structure will allow us to aggregate occupations into the proper classes within which mobility is easy and typical, and thereby to ground our class scheme solidly in the problem of mobility.

In this paper we conceptualise the mobility table as a network of occupations tied together by the social mobility between the occupational categories. We develop an exploratory method that clusters together the occupations in order to derive the intra-generational social class structure. In the following section we elaborate on the intrinsic relationship between class and the structure of mobility in the tradition of class-analysis. We argue that, in order to advance class-analysis, we need to approach the empirical analysis of classes and their boundaries in a descriptive way, as opposed to a purely explanatory strategy based on statistical model-testing. We then move on to the main contribution, namely, the new analytical method we propose. Subsequently, we apply the method to Danish register-data and analyse 1,590,834 job shifts which took place between 2001 and 2007. After presenting the results, we discuss certain methodological issues and the resulting social class categories in relation to the factors of gender, income, education, and political action. Finally, we present our conclusions.

2. Class and mobility structure

For students of social mobility, Weber's above-quoted definition of a social class constitutes an important starting point. Weber makes two points: First, intra- as well as inter-generational mobility is constitutive of social class. Inter-generational mobility has received much attention, but in this paper we are going to analyse intra-generational mobility. This should suffice, because the purpose of this paper is to present and demonstrate our method's ability to overcome the problem of identifying class boundaries. Furthermore, intra-generational mobility has received relatively little attention in the literature on class, and this study contributes to countering this bias. It does imply, however, that the results should not be directly compared to analyses of inter-generational mobility. Different dynamics are at play in the two forms of mobility, which warrants separating the analyses of the two forms of mobility and their class structure. We therefore also hesitate to make direct comparisons to class schemes like those mentioned above, which, even though they may claim to also encompass intra-generational mobility, are often constructed in relation to inter-generational mobility.

Weber's second, and more crucial point is that social class is made up of patterns of "easy and typical" mobility. Theoretically speaking, a mobility pattern refers to the transition from one position in the economy to another. When operationalised empirically, mobility pattern typically means movement from one occupation to another. The logic of Weber's definition further suggests that analysis of mobility patterns should be the starting point for the empirical investigation of the social class structure.

Giddens rephrases and specifies Weber's definition of social classes as '[…] a cluster of class situations which are linked together by virtue of the fact that they involve common mobility chances, either within the career of individuals or across the generations.' (1973: 48). What we develop in this paper is precisely a method, which can identify such "clusters of class situations" based on the observed patterns of mobility at a highly disaggregated level. Also, following Giddens, we separate the question of intra- and inter-generational mobility, and focus on the former. This represents a deviation from Weber who tended to conflate the two forms of mobility. An even more operationalized definition, alluding to social network analysis (SNA), can be found in Breiger's summary of Blau, Duncan and Giddens' Weberian perspective: '[…] classes are essentially sociometric "cliques" defined so that mobility chances […] are higher *within* the cliques than *between* them' (Breiger, 1981: 584).

The logic of Weber's definition of social class suggests an approach in which description is prior to explanation. In order to investigate the mechanisms generating boundaries to mobility, i.e. explain social classes, we must first identify the exact location of the social class boundaries. To use the concepts of Giddens, we must carefully describe the class based structure of social mobility in order to investigate the structuration of social mobility. Following this logic, Giddens underscores that '[…] the existence of distinct class 'boundaries' […]' cannot be '[…] settled *in abstracto*', because the class structuration of various societies '[…] differs significantly according to variations in economic and political development' (Giddens, 1973: 110).

Hence, what is needed is a detailed description of the social class structure, understood as a number of occupational clusters within which mobility is relatively high and between which mobility is relatively low. Such a description will provide the structure that is to be explained. The explanatory effort should take its point of departure in investigation of the identified boundaries. This would enable examining the mechanisms and processes generating the barriers to mobility.

An argument supporting this claim has been provided by Savage et al. (2013) who, echoing Giddens, argue that the need for descriptive and exploratory methods has been accentuated by the growing realization that no single class scheme fits all societies. Rather, if uncritically applied such universal class schemes may be misleading because they make us blind to significant differences between societies, and explanations related to specific societies may be overlooked. This is due to both '[…] real cross-nation differences with respect to qualification levels, job autonomy, career prospects (i.e. social mobility), organization of production, etc.' (Savage et al., 2013: 223) and the multi-dimensionality of social stratification. In our perspective, multi-dimensionality means that the mobility barriers generating the social mobility structure may consist of a variety of processes and mechanisms of selection. This implies that, rather than excluding the impact of factors such as gender, age, institutions, cultural norms, ethnicity, etc. on mobility, they are perceived as potential drivers of class formation.

For instance, if a social class turns out to be the result of mobility patterns created by young people enrolled in education who work part-time jobs in the service sector and often change jobs, this is not a problem, but a result in the sense that the existence of this social class is explained by a specific position in the economy tied to age. An example of such a finding is provided in the case of the *Emergent service workers* identified by Savage et al. (2013: 240ff.). Following this logic, also such factors as e.g. gender-based division of labour may give rise to barriers to job-mobility constitutive of social classes. Such a descriptive and exploratory approach implies that, in the subsequent mapping of the Danish intra-generational social class structure, we do not, a priori, exclude or seek to control for any number of factors that according to a given theory of class are considered extra-class.

The descriptive concept of social class does, of course, relate to class theory, but the distinction is crucial as the concept of social class does not imply a theory of its formation as opposed to the class-concept of a class theory. Thus, several social classes may be combined into one class due to theoretical explanation. For instance, if unskilled workers are separated into different social classes the separate social classes may together constitute a class of unskilled workers. Thus, social classes should be viewed as the building blocks or segments that constitute the starting point for the operationalization of a class theory into a class scheme. However, the social classes impose limitations on the construction of classes. The salience of mobility barriers to class boundaries implies that if a class theory suggests splitting a social class (within which mobility is easy and typical) into different classes, the validity of the class scheme will be undermined. This is due to the permeability of the class boundaries resulting from the split.

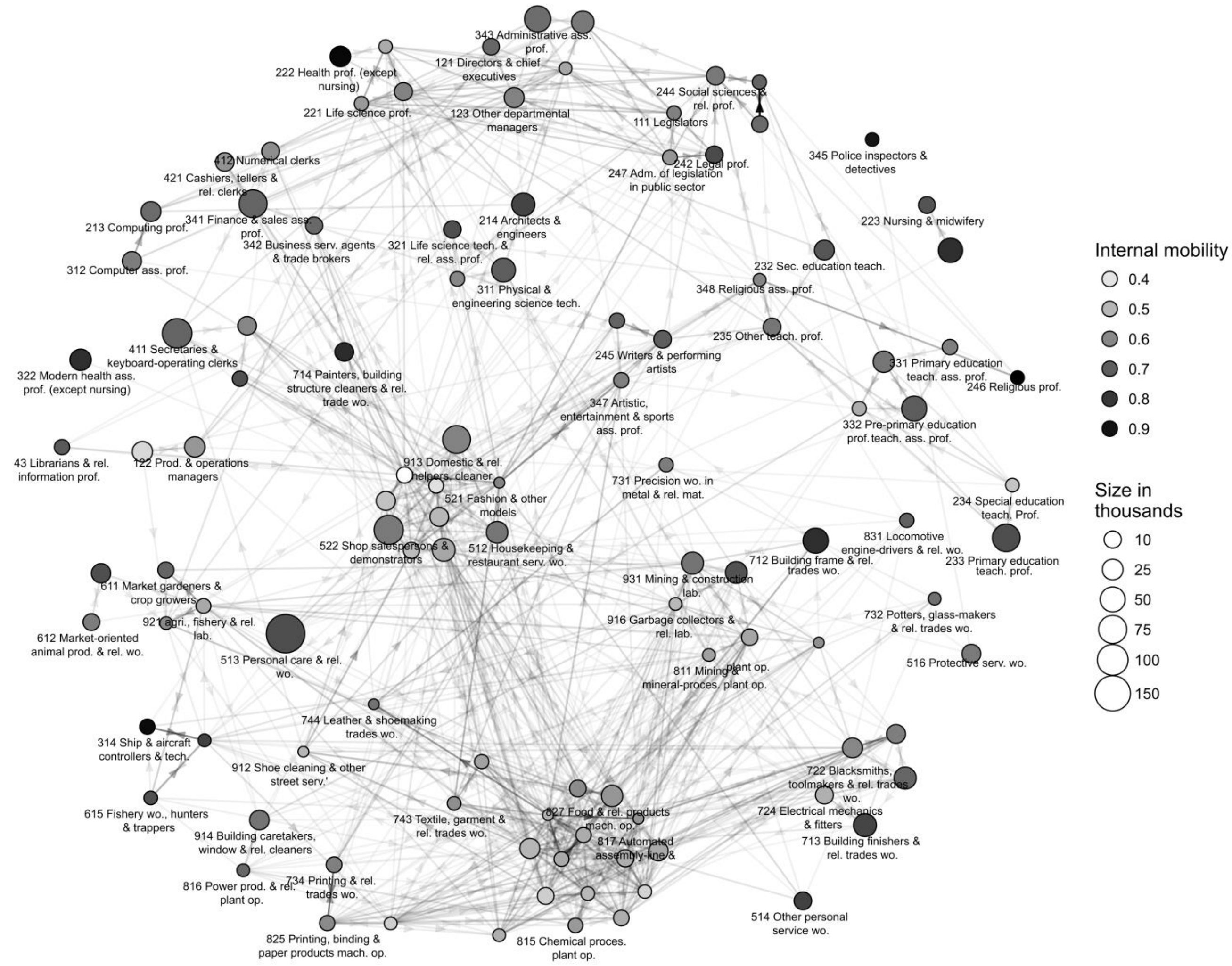
3. The mobility table as network

This study relates to the inductive tradition of research on class and stratification with the occupational mobility table at the centre (Breiger, 1981; Goodman, 1981; Klatzky and Hodge, 1971; Levine, 1972; MacDonald, 1972). However, this study diverts in two ways. First, we focus on intra- and not inter-generational mobility. Second, most scholars in this tradition have aimed at identifying the dimensions driving mobility and have often based their models on scales of the occupational hierarchies (E.g. Duncan, 1961; Treiman, 1977). This study instead aims at describing the intra-generational social class structure in an exploratory manner.

In order to correctly delimit the mobility categories in a descriptive way, we conceptualize the occupational mobility table as a network of occupations. This has been done before by e.g. Grifftihs and Lambert (2012), although they analysed marriage-relations and not social mobility. Occupations take the form of nodes in the network and the individuals' mobility between occupations generates the ties. From such a mapping we can observe between which occupations the labour force flows freely and between which occupations barriers appear to disrupt the flow of labour. The next step is quite literally to identify the '[…] essentially sociometric "cliques" […]' (Breiger, 1981: 584) that constitute the social class categories of the mobility structure.

The ties connecting the nodes are measured as the relative risk (RR). RR expresses the relative likeliness of the occurrence of job shifts from one occupation to another. RR=1 represents the ideal situation of 'perfect' mobility. The straight forward translation of the RR into a measure of the intensity (or weight) of a network tie would be that if RR≥1 there is a tie between the nodes. If RR<1 the nodes are not connected. We can now depict the intra-generational mobility on the Danish labour market in the period 2001-2007 as a directed weighted network which is done in figure 1. Later we present the data used to construct the network. The size of the nodes expresses a logarithmic function of the size of the occupational category. The colour of the nodes expresses degree of internal mobility. The darker the node, the closer to 100% internal mobility. The alpha of the colour of the edges expresses the intensity of the mobility flow.

Figure 1. Intra-generational occupational mobility as network



4. The Mobility Network Clustering Algorithm (MONECA)[2]

In a dense network like the one depicted in figure 1, in which almost all nodes are connected, it is difficult to make sense of the structure. In this case we look for groups of especially tightly connected nodes. The aim is to identify the cohesive and non-overlapping sub-groups of the network in order to derive the social class structure. Conventional SNA concepts such as clique and core are problematic as they produce cluster solutions that overlap. In contrast, clusters generally refer to *non*-overlapping cohesive sub-groups (Scott, 2000: 126ff.). Cluster analysis has the additional advantage over the sociometric concepts that it is better suited to handle weighted networks (Knoke and Yang, 2008: 80ff.). Others have used cluster analysis on distance-matrices of social mobility (e.g. Hope, 1972). The novelty of this study is the development of a cluster algorithm that clusters on the basis of weighted network ties rather than abstract distances. The principal difference is that traditional cluster analysis is based in the abstract space of the distance matrix where all categories, no matter how remote, have a distance and therefore, in principle, can be clustered. The binary logic of network ties, on the other hand, implies that unconnected categories cannot be clustered together (see Toubøl et al. 2013 for a detailed discussion of this issue).

MONECA is designed to identify discrete clusters of interconnected nodes in a dense network. The logic of the algorithm is closely associated with the concept of the clique. The task of the algorithm is to decide to which clique to allocate the nodes that are in the overlapping area. To answer this question, the algorithm works in an agglomerative manner, considering the connections of the single pairs of nodes. The first step is to pair together the two most intensely connected nodes, which then form a cluster with the properties of a dyadic clique. In step two it proceeds to pair together the second most intensely connected pair of nodes in the same manner as in step one. And so it continues until all connections have been considered. At any subsequent step, if two nodes under consideration are already members of two different clusters, they can only be joined together if all the nodes in their respective clusters constitute a clique and, thus, can be joined together forming a new cluster. This clique-criterion provides the stop rule for when no more single or sets of nodes should be paired together forming new clusters (for an expanded explanation of the methodology, see appendix A).

5. The Danish intra-generational mobility structure

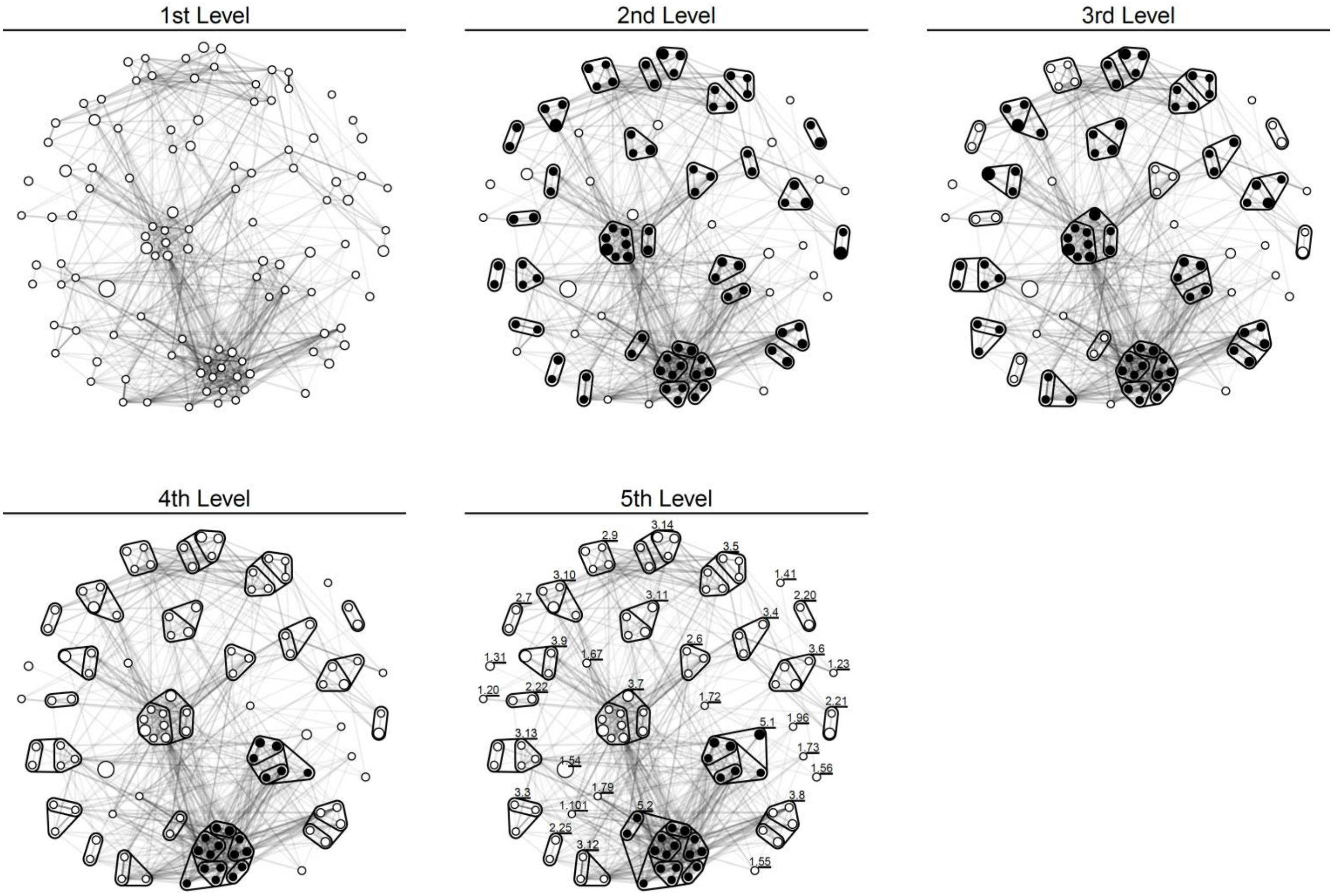
The application of the above outlined approach to Danish labour market mobility data covering the period 2001-2007 demonstrates how a network analytical approach may provide a solution to the longstanding problem of identifying class boundaries.

The data of our study was collected by Statistics Denmark and comprises the entire Danish labour market in the period 2001-2007[3]. This provides us with yearly information concerning the individuals' occupations (ISCO) as well as information on job shifts. With this information we can construct career sequences from 2001-2007 consisting of up to seven states in the cases of individuals employed during the entire period. We can determine whether the individuals changed their job or not as well as from, and to which occupation they moved during the transition from one state to another. In total, 11,274,435 transitions are recorded throughout the period. Of these, 1,590,834 (14.1%) represent job shift transitions. We disaggregate the occupational coding to the three-digit level of ISCO leaving us with 109 occupational categories[4] and develop a 109 x 109 occupational mobility table with 11,881 cells.

For the calculation of the RR, we weight the expected frequencies by the proportion of the total number of employees in the occupational category. To be precise, we calculate the expected frequencies of job shifts as the mean of the column and row proportions of the total number of transitions to the grand total of job-shift-transitions. Thus, the expected frequency of job shifts in a given cell depends on the total size of the row and column cell of the occupation. The reason for doing so is to avoid underestimating occupations with a large proportion of mobile employees, as well as overestimating occupations with a small proportion of mobile employees. Furthermore, despite the large amount of data, some cells in the mobility table are very sparsely populated. In order to avoid false connections due to measurement errors, cells with a frequency below the threshold of five have been erased.

The clusters are identified as described in section 4, meaning they are cliques in which all nodes are mutually connected, i.e. density=1. The criterion that the mobility pattern has to be mutual in order to constitute a tie is important, as we would otherwise risk joining together occupations which are linked through promotion, such as junior and senior positions in a hierarchical organisation. Thus, one-way mobility patterns do not constitute ties considered by MONECA. Forming new categories from the identified clusters, the number of categories is reduced from 109 at the original level 1 to 56 at level 2. MONECA can be used to reduce the number of categories further. We can form a new 56 x 56 directed matrix using the categories identified by MONECA at level two and perform the procedure once more. The result of what constitutes a third level in this agglomerative clustering procedure is shown in the level 3 graph of figure 2. The new clusters are marked by lines encircling existing clusters of nodes and single nodes. The new clusters are, by definition, not cliques but are, nonetheless, very dense.

Figure 2. The five levels of the cluster analysis

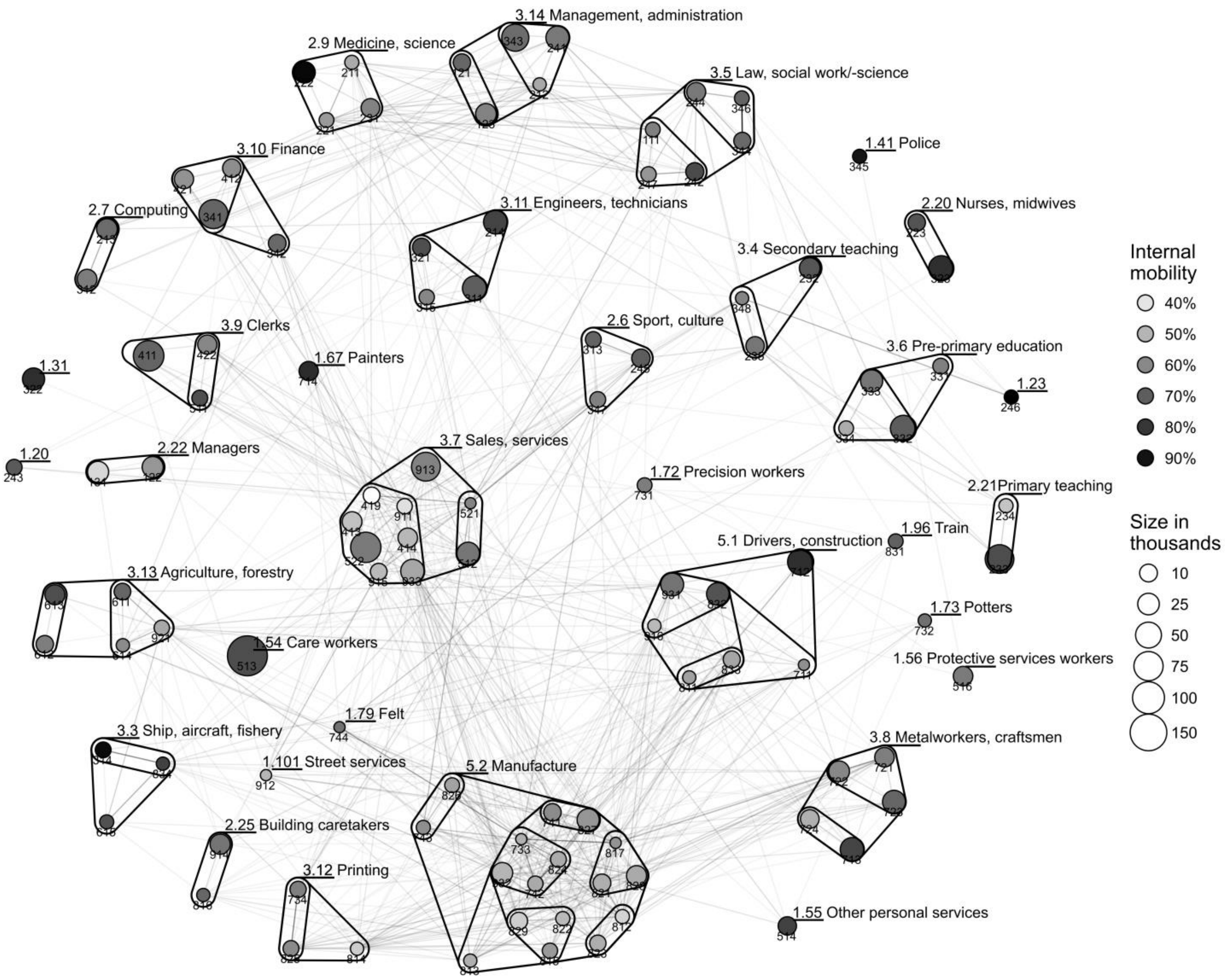


We can repeat the procedure a number of times, reducing the number of categories further (see figure 2). The amount of possible repetitions is restricted because the matrix is not hollow. As such, for each repetition which results in more aggregated categories, a larger share of the total mobility will be within-mobility located in the diagonal. This leaves less between-mobility to constitute mutual mobility patterns. Eventually, no mutual connections can be detected; hence, no more categories can be merged. Thus, MONECA has a built in stop point that is conditioned by the chosen cut point (in this case the cut point is RR=1). Thus, in contrast to most agglomerative clustering algorithms, MONECA does not continue until all cases are merged into one big cluster. In this case, MONECA continues until level 5, giving us 34 categories, as can be seen from the level 5 graph of figure 2.

6. Results

The result is a cluster solution of 34 occupational categories, as depicted in figure 3. Two categories are level five clusters, 12 are level three clusters, seven are level two clusters and 13 are level one clusters (for details regarding the levels of the cluster solution see appendix B).

Figure 3. Result of cluster analysis



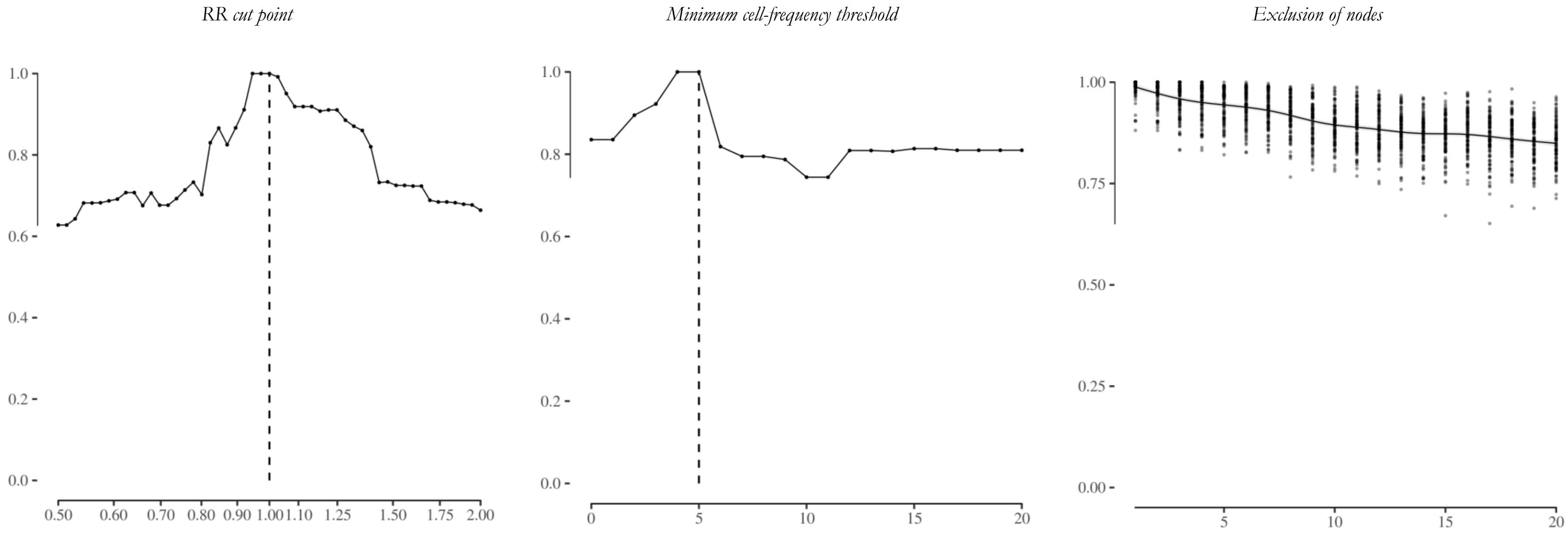
A cluster solution should always be critically inspected and assessed by the researcher. In our case this is especially important with regard to clusters formed at level three and higher. By virtue of the clustering procedure these are not cliques where movement between all occupations is easy. Hence, a critical inspection and evaluation of whether the higher level clusters are too incoherent or their boundaries are too permeable is paramount. Summary statistics of the clusters are provided in table 1. Numbers correspond to the numbers in figure 3. Titles are the product of our interpretation of what characterises the occupations of the clusters. Size refers to the share of the total number of employees at the labour market. The 10 largest categories cover 78% of the labour market. The concentration in size suggests a relatively simple class structure with a few meso-classes and several micro-classes.

Table 1. Summary of final cluster solution produced by MONECA by size

#	*Title*	*Size*	*Density*	*Within mob.*	*Nodes*
3.7	Sales, services	0.205	0.811	0.750	10
1.54	Care workers	0.111	n/a	0.740	1
5.1	Drivers, construction	0.079	0.476	0.812	7
5.2	Manufacture	0.066	0.603	0.673	17
3.8	Metalworkers, craftsmen	0.066	0.850	0.802	5
3.14	Management, administration	0.058	0.650	0.723	5
3.9	Clerks	0.055	0.667	0.701	3
3.1	Finance	0.054	0.667	0.713	4
3.6	Pre-primary education	0.043	0.833	0.818	4
3.11	Engineers, technicians	0.042	0.833	0.790	4
2.21	Primary teaching	0.027	1.000	0.750	2
2.20	Nurses, midwives	0.024	1.000	0.860	2
3.5	Law, social work/-science	0.023	0.600	0.786	6
2.9	Medicine, science	0.021	1.000	0.848	4
2.7	Computing	0.020	1.000	0.770	2
1.31	Health associate professionals	0.014	n/a	0.820	1
3.4	Secondary teaching	0.013	0.667	0.705	3
2.22	Managers	0.013	1.000	0.617	2
2.6	Sport, culture	0.012	1.000	0.727	3
3.13	Agriculture, forestry	0.011	0.600	0.730	5
2.25	Building caretakers	0.009	1.000	0.652	2
1.67	Painters	0.009	n/a	0.829	1
3.3	Ship, aircraft, fishery	0.006	0.667	0.901	3
1.55	Other personal services	0.006	n/a	0.771	1
3.12	Printing	0.005	0.667	0.680	3
1.56	Protective services workers	0.004	n/a	0.632	1
1.20	Archivists	0.001	n/a	0.709	1
1.23	Religious professionals	0.001	n/a	0.935	1
1.72	Precision workers	0.001	n/a	0.625	1
1.96	Train	0.001	n/a	0.694	1
1.41	Police	0.000	n/a	0.893	1
1.73	Potters	0.000	n/a	0.657	1
1.79	Pelt and leather workers	0.000	n/a	0.653	1
1.101	Street services	0.000	n/a	0.500	1

Density is the share of the total number of possible edges in the cluster; the higher the density, the more coherent the cluster is, because the barriers to mobility within the cluster are low. The ideal is a clique where density equals one. All level one and two clusters are by virtue of MONECA cliques. The opposite is the case for higher level clusters, as they never have a density of one. Still, with the exception of the level five cluster *5.1 Motor vehicle drivers & construction* the density for all clusters is high (≥ 0.6). Within-mobility is the share of the mobility from a given cluster going to a destination within the same cluster. This indicates the discreteness of the clusters: high within-mobility means that a cluster is relatively isolated in the social mobility structure in the sense that the chance of entering or leaving the cluster, i.e. crossing the social class boundary, is small.

Figure 4. Reliability test results: correlations with final solution



Inspecting table 1, the relatively low density of cluster 5.1 draws attention. Close inspection of the map shows that the latest addition, *712 Building frame & related trades workers*, is only connected to two of the other six occupations in the cluster, meaning that mobility from this occupation is not easy. Depending on subject matter, it might be advisable to break up the cluster and return to the level four solution, but for the sake of simplicity, and because the purpose of this paper is to demonstrate the workings of the method, we stick with MONECA's original cluster solution.

Before we turn to the discussion, we will provide the main conclusion from three reliability tests. The tests are presented in detail in appendix C. The tests concern the RR cut point, the minimum cell-frequency threshold, and sensitivity to changes in the network. Figure 4 shows the correlation between the final cluster solution and the manipulated solutions of the three tests.

The cut point of RR=1 is qualified on theoretical grounds because it is equal to perfect mobility. Nonetheless, it would be critical to the reliability if small variations cause massive change to the cluster solution. This is not the case. When comparing the cluster solutions of different RR cut points to the solution of an RR=1, the correlation never goes below 0.6 within the entire test-interval of $0.5 \leq RR \leq 2.0$, and within the $0.825 \leq RR \leq 1.3$ it is consistently higher than 0.8. The minimum cell-frequency threshold is more critical as its determination to an extent relies on the judgement of the researcher. However, the test-results are reassuring. The correlation with the solution of a threshold of five remains around 0.8 in the entire test-interval from 0 to 20. Finally, we test MONECA's ability to reproduce the results when erasing up to 20 randomly chosen nodes in the network, which, in turn, affects the relations between the remaining nodes. We compared the results of the reduced networks with the solution of the complete network. This was done 100 times at each level from 1-20. On average, the correlation was never lower than 0.9 and the single lowest correlation among the 2000 was 0.65. From this we conclude that MONECA, in a fairly consistent manner, reproduces the results even when imposing rather comprehensive changes to the network.

7. Discussion

We begin by considering the sociological status of the categories suggested by MONECA. We then discuss the main limitations of the proposed methodology. Finally, we go beyond the dimension of mobility and consider the correspondence between the categories provided by MONECA and the factors of income, education, political behaviour and gender which represent key analytical aspects of class and stratification. In what follows, we only consider class in relation to intra-generational mobility. We do not include the class structure of inter-generational mobility, as we consider this to be a separate question.

In most studies, 34 class categories are neither analytically relevant, nor practical. As argued earlier, the 34 intra-generational social classes can be aggregated into fewer classes according to, for instance, a certain hierarchical principle, such as education or income, or other theoretical considerations. Mobility within these aggregated classes would naturally not be as likely as within the intra-generational social class categories. However, the barriers between the aggregated classes would remain in place and we are not in danger of creating a class scheme with permeable boundaries.

For theoretically derived class schemes this is not the case. As they are not derived from data on social mobility patterns they will quite likely place occupational categories with strong mutual mobility flows in separate classes, i.e. split social classes. This problem is compounded if you impose a class scheme, designed to fit a specific country in a specific historical context on another country in another historical context.

Thus, we argue that the 34 occupational clusters should be seen as intra-generational social classes which, by themselves, are of analytical interest as well as possible building blocks, from which a simpler class-map can be formed. For instance, through an empirical analysis of the categories' relation to factors such as property-relations, life-chances or social interaction we may find that two or more of the categories share the same position in the hierarchy and, thus, are segments in the same class. The result would then not violate the mobility criterion.

The method we have developed and presented in this article has two important interrelated limitations. First, the occupational classification may weaken the validity of the results if the categories do not fit the actual job-positions in the economy, i.e. are invalid. Due to such misfits, the occupational classifications may group together dissimilar jobs or divide similar jobs into different categories. The latter scenario ought not to be a problem to the approach of this paper, as two categories with similar jobs implying a high level of inter-categorical mobility would be clustered together. The former scenario, in which dissimilar jobs are grouped together in the same category, is harder to solve. In fact, we cannot be sure that this is not the case. However, the more disaggregated and detailed the baseline occupational classification, the lower the chance that dissimilar jobs are grouped in the same category.

Second, the number of observations in the data imposes limits to how disaggregated a level we can start from. A too sparsely populated mobility table lowers the validity of the subsequent analysis. The methodological approach of this paper, designed to start from a very disaggregated level in order to partially avoid the mentioned hazards inherent in any occupational classification, depends on access to large datasets on mobility. Even though few countries collect register data covering the entire population, as Denmark does, digitalization will most likely result in more such datasets becoming available to researchers in the future. Furthermore, large national survey programs provide promising sources of such large datasets on mobility by pooling the data from several rounds.

While describing the patterns of intra-generational mobility and -social class segregation in Danish society is interesting in itself, it is crucial to show that these social class categories are of importance to more than mobility. Figure 5 shows the occupations measured by *gender* (proportion of women), *income* (the crude mean of the employees' total disposable income per year after taxes, benefits and payment of interests and alimony from tax-records in DKK. 2007 DKK/£ exchange rate≈11), *union density* (proportion of employees who were member of a trade union) and *education* (mean number of prescribed years of education of the highest level of education received). In table 2, the same variables are presented by social class in descending order by income (for the exact numbers by occupation see appendix D and for statistics summarizing the distributions on income and education see appendix E).

Starting with income, we find that, in general, the occupations within the same social class are in the same income-range. When variation occurs, as in the case of classes *3.13 Agriculture, forestry*, *3.14 Management, administration* and *3.5 law, social work/-science*, this variation is reflected by the subgroups of the social class.

Turning to the between variation we find large differences when we inspect the mean disposable incomes in table 2 which is ordered by income (descending). At the very top of the income scale *2.22 Managers* distinguishes itself from the rest with a disposable income of DKK573,365, more than four times that of *3.7 Sales, services* (DKK134,150) at the bottom of the table and the income scale. This exclusive management group only makes up 1.3% of the labour market[5]. Next follows a group of nine high income social classes, in descending order from *3.3 Ship, aircraft, fishery* (DKK324,543) to *3.1 Finance* (DKK256,265). They make up almost a quarter (23.7%) of the labour market,. Then follows a large group going from *1.96 Train* (DKK231,657) to *1.101 Street services* (DKK178,708). This group constitutes 42.8% of the labour market. Finally, we have a group of four relatively low income social classes going from *Potters* (DKK170,028) to *3.7 Sales, services* (DKK134,150). Together they make up almost a third of the labour market, 32.2%.

Following these observations, a classification based on a hierarchy of income suggests four aggregate classes: 1) A tiny elite class consisting of top level management, 2) an upper middle class consisting of lower level management, academics, administrators, specialists and the financial sector, 3) a large and heterogeneous middle class consisting of both blue and white collar workers, and 4) a large lower class consisting of service workers.

Figure 5. Occupations by income, education, union density and gender distribution in 2007

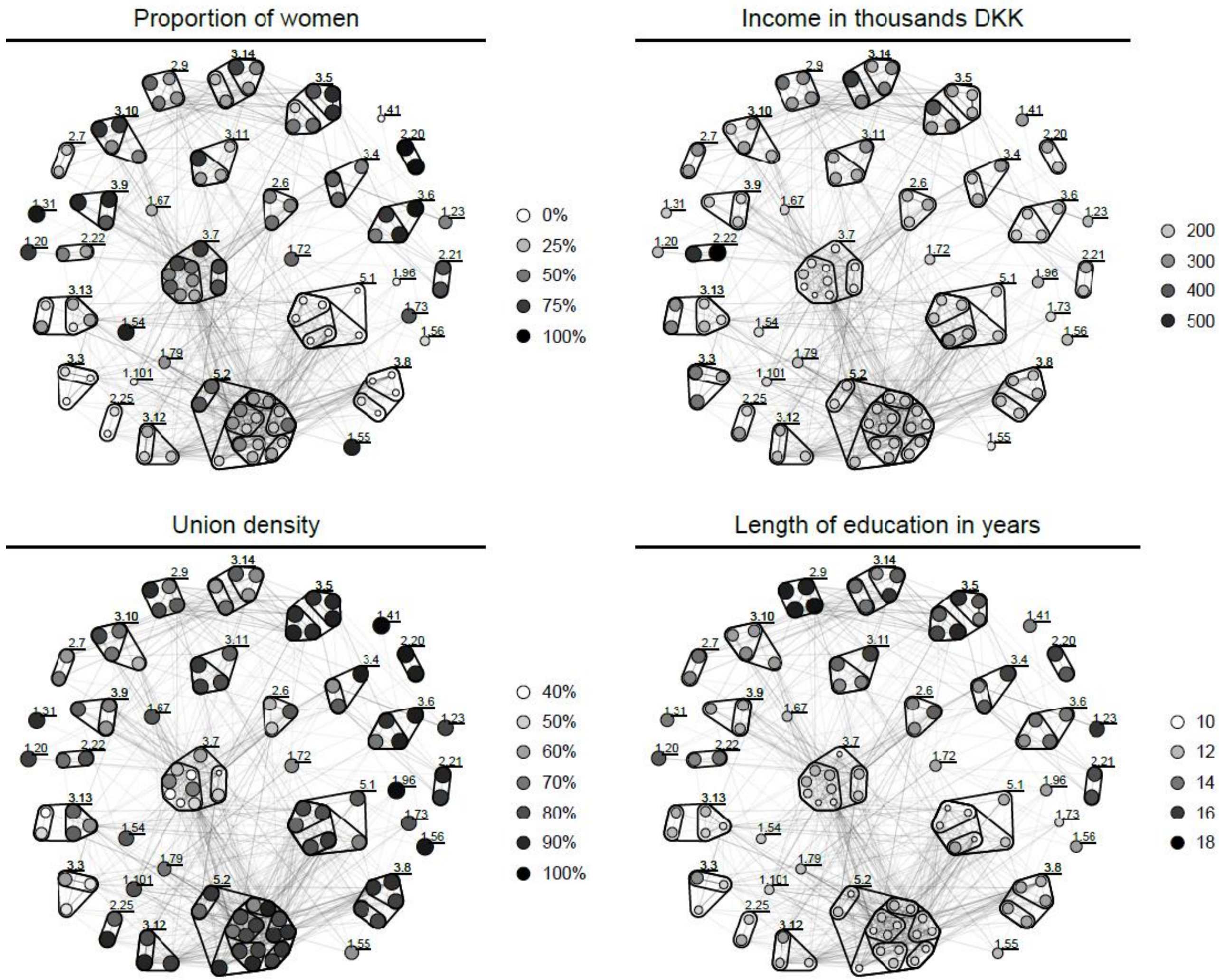


Table 2. Clusters by income, education, union density and gender distribution in 2007

#	*Disp. Income* (DKK)	*Education, years*	*Union density*	*Share of women*
2.22	573,365	14.1	77.6%	33.3%
3.3	324,543	12.8*	57.3%	9.4%
2.9	319,998	17.2	82.6%	47.9%
2.7	283,378	14.0	64.5%	20.7%
1.41	274,551	13.6	98.6%	2.9%
3.11	263,466	14.5	78.5%	31.3%
3.14	260,838	13.6	72.5%	62.3%
3.5	259,903	15.7	89.2%	64.6%
3.4	257,314	15.3	86.0%	48.9%
3.1	256,265	13.2	70.2%	48.3%
1.96	231,657	12.7	97.9%	4.5%
3.13	231,378	11.9	59.5%	23.1%
2.6	229,191	14.1	66.9%	45.6%
1.23	224,456	16.3	82.2%	44.7%
1.56	222,799	12.7	95.2%	13.6%
3.12	219,583	12.0	85.9%	21.5%
2.20	218,122	15.1	93.2%	95.4%
2.21	216,611	15.3	85.6%	66.5%
1.20	215,744	14.9	83.6%	75.1%
1.79	206,694	11.8	70.5%	38.8%
3.8	203,340	12.5	83.8%	2.6%
5.1	195,615	11.3	78.7%	3.7%
3.9	192,325	12.6	72.6%	82.7%
1.72	192,225	12.5	65.1%	55.4%
3.6	191,953	14.2	89.6%	81.5%
1.31	191,830	13.8	88.7%	91.3%
2.25	189,132	11.6	75.5%	10.2%
1.67	181,745	11.8	77.7%	26.9%
5.2	181,414	11.1	82.3%	35.3%
1.101	178,708	11.6	78.7%	4.3%
1.73	170,028	11.3	77.7%	63.6%
1.54	164,446	11.6	77.0%	87.8%
1.55	145,392	12.0	63.3%	86.3%
3.7	134,150	11.1	50.7%	55.9%
Total	201,561	12.7	74.1%	51.9%

* The very low mean years of education is due to aircraft pilots and air traffic controllers being registered with unrealistically short educations for reasons unknown (Albæk and Thomsen, 2011: 28).

Such a simple four-group hierarchy becomes problematic if we take education into consideration. The correspondence between level of income and education is far from perfect. For instance, the members of group *2.22 Managers* with a mean income of DKK573,365 have a mean of 14.2 years of education. Members of *3.6 Pre-primary education* who also have a mean of 14.2 years of education, on the other hand, have an income of less than half that amount (DKK191,953). At the bottom of table 2, people in low income category of *3.7 Sales, services* unsurprisingly also have relatively short educations (DKK134,150/11.1). However, surprisingly, people in the groups *5.1 Drivers, construction* and *5.2 Manufacture,* who have equivalent levels of education, have a much higher income (respectively DKK195,615/11.3 and DKK181,414/11.1). These discrepancies between income and education indicate that the individuals' income is not only a product of human capital possession but also depends on their (social) class membership.

The relationship between class and political action is central to the discussion of the relevance of class. Usually, this has been discussed in terms of political partisanship. We take a slightly different approach and look at the degree of unionization. Even though unions are not political parties, they are political organizations and in Denmark, unions have historically been strongly associated with the political left (Toubøl and Jensen 2014).

In international comparison, the overall union density in Denmark is, at 74.1%, very high. Despite the general high level of unionization, there are considerable variations in union density among occupations. This may indicate that unionization is one of the factors influencing income distribution. An example of this is the contrast between *3.7 Sales, services,* which has the lowest union density of all social classes (50.7%,) and *5.1 Drivers, construction* and *5.2 Manufacture* with union densities well above average (78.7% and 82.3% respectively). A possible explanation of why the income level relative to educational level is so much higher in 5.1 and 5.2 compared to 3.7 is that unions are much stronger in 5.1 and 5.2. However, unions do not explain the low income level relative to the level of education of *3.6 Pre-primary education*, *2.20 Nurses, midwifes* and *2.21 Primary teaching.* These categories have a high union density, well above average, but they still have low incomes relative to the level of education. We suggest that the explanation may be related to the extra-class factor of gender.

Gender is related to stratification theory but it is usually treated as a dimension distinct from class. However, the growing awareness of the intersectional nature of inequality (e.g. Walby et al., 2012) calls for the combination of class and gender, among other factors. In figure 5 and table 2 we register great variation in the proportion of women among occupations. The intra-generational social class categories reflect these variations. The low-income category *3.7 Sales, services* has a relatively high proportion of women (55.9%) compared to *5.2 Manufacture* (35.3%) and especially *5.1 Drivers, construction* (3.7%). However, we find the highest proportion of women among all social classes in *2.20 Nurses, midwifes* (95.4%), while also *3.6 Pre-primary education* (81.5%) and *2.21 Primary teaching* (66.5%) have proportions of women well above the overall average of 51.9%. These correlations indicate a negative relationship between income and the proportion of women in a social class.

8. Conclusion

The preceding discussion argues that employing descriptive methods like MONECA makes it possible to investigate the multidimensional nature of class structuration. Hence, the above outlined method and approach call for a more nuanced theoretical approach to class. The results also underline the need to focus on society and context specific explorations of class structure, rather than universal class schemes. The mapping of the Danish intra-generational social class structure 2001-2007 in this article suggests that things are much more complex than what any single universal class scheme and theory is able to grasp. Furthermore, given this complexity, it is only reasonable to expect significant variation between different societies.

The above outlined methodological approach provides two innovations directed to the task of society and context specific analysis of class structure. First, the conceptualization of the mobility table as a network enables us to depict the social mobility structure in a transparent and intuitively meaningful way, which corresponds with the basic imagery of social class and social mobility theory. Based on a simple probability model of randomness, mobility barriers become visible. Furthermore, we can distinguish between the direction and the relative intensity of the flow and take the differences in proportion of mobile workforce and the proportion of internal mobility into account.

Second, MONECA enables us to identify discrete clusters of occupations between which mobility is easy and typical, and it provides a methodological means to give substance to Weber's rudimentary definition of social class. The clusters identified provide the basic intra-generational social class categories. These should be perceived as the result of an exploratory and descriptive analysis, mapping the social class categories of the intra-generational social mobility structure. However, subsequent explanatory analysis should have these categories as a basic starting point. The social classes may be aggregated into fewer and larger classes if an extended definition of class is employed. However, a social class category should never be split into different categories within a class scheme as the boundaries of the new class categories would be permeable. The validity of such a class scheme would not meet the mobility criterion.

As indicated by the preliminary and strictly descriptive analysis of the prominent variables income, education, gender and political action, the intra-generational social class map identified, reflects important factors of differentiation. Through the precise identification of boundaries, the social class map derived in the manner outlined above may very well provide a better starting point for the investigation of the factors driving class formation than the hitherto dominant theory-driven approaches.

Going beyond the question of class and social mobility, MONECA is useful for the analysis of any dense, weighted network from which it is meaningful to extract discrete categories based on a principle of connectivity. Thus, the outlined exploratory and data driven approach to categorization may prove very useful in providing the fundamental categories for research into other questions than social mobility.

Acknowledgements
We are grateful to Thomas P. Boje, Peter Gundelach, Bue Rübner Hansen, Christian Lyhne Ibsen, Carsten Strøby Jensen, and Jens-Peter Thomsen as well as the anonymous reviewers for indispensable comments and suggestions.

Notes
1. The data-package MONECA written in R is freely available from https://github.com/antongrau/moneca. It includes tools for transforming a table of associations into a weighted network and subsequent analysis with the MONECA algorithm including outputs summarising the process as well as providing descriptive measures useful for assessing the qualities of the clusters, some of which has been discussed in this article and provided in the appendices.

2. The argument of this section have previously been presented in a modified version at the XXXIII Sunbelt Social Networks Conference (Toubøl et al., 2013).

3. This definition excludes persons who are not active on the labour market. The most notable excluded groups are children, unemployed, persons on leave including sick leave or parental leave, senior citizens and affluent persons living from capital returns.

4. The Military has been excluded because this occupational category covers a very diverse range of jobs and qualifications, which makes it analytically troublesome. Danish military is relatively small and the exclusion is not as problematic as in the cases of countries with large militaries.

5. However, even within this group, an even more exclusive income group exists, which is revealed by comparing the mean of DKK573,365 to the median of DKK264,893 (See appendix E).

References

Albæk K and Thomsen LB (2011) *Er Kvindefag lavtlønsfag? En analyse af sammenhængen mellem løn og andelen af kvinder i enkelte arbejdsfunktioner*. SFI-rapporter, København: SFI--Det nationale forskningscenter for velfærd.

Andersen PL and Hansen MN (2012) Class and Cultural Capital—The Case of Class Inequality in Educational Performance. *European Sociological Review* 28(5): 607–621.

Blau PM and Duncan OD (1967) *The American Occupational Structure*. New York: John Wiley & Sons.

Bourdieu P (1984) *Distinction, a social critique of the judgement of taste*. Routledge & Kegan Paul.

Bourdieu P (1987) What Makes a Social Class? On The Theoretical and Practical Existence Of Groups. *Berkeley Journal of Sociology* 32: 1–17.

Breiger RL (1981) The Social Class Structure of Occupational Mobility. *American Journal of Sociology* 87(3): 578–611.

Clogg CC and Goodman LA (1984) Latent Structure Analysis of a Set of Multidimensional Contingency Tables. *Journal of the American Statistical Association* 79(388): 762–771.

Duncan OD (1961) A Socioeconomic Index for all Occupations. In: Reiss AJ (ed.), *Occupations and social status*, New York: Free Press of Glencoe, pp. 109–138.

Erikson R and Goldthorpe JH (1993) *The constant flux, a study of class mobility in industrial societies*. Clarendon.

Erikson R, Goldthorpe JH and Hällsten M (2012) No way back up from ratcheting down? A critique of the "microclass" approach to the analysis of social mobility. *Acta Sociologica* 55(3): 211–229.

Giddens A (1973) *The class structure of the advanced societies*. London: Hutchinson.

Goldthorpe JH (2007) *On sociology*. 2. ed. Stanford University Press.

Goodman LA (1981) Criteria for Determining Whether Certain Categories in a Cross-Classification Table Should Be Combined, with Special Reference to Occupational Categories in an Occupational Mobility Table. *American Journal of Sociology* 87(3): 612–650.

Griffiths D and Lambert P (2012) Dimensions and Boundaries: Comparative Analysis of Occupational Structures Using Social Network and Social Interaction Distance Analysis. *Sociological Research Online* 17(2): 5.

Grusky DB and Weeden KA (2002) Class Analysis and the Heavy Weight of Convention. *Acta Sociologica* 45(3): 229–236.

Hope K (1972) Quantifying Constraints on Social Mobility: The Latent Hierarchies of a Contingency Table. In: Hope K (ed.), *The Analysis of social mobility; methods and approaches*, Oxford studies in social mobility. Working papers, Oxford: Clarendon Press, pp. 121–190.

Jonsson JO, Grusky DB, Carlo MD, et al. (2009) Microclass Mobility: Social Reproduction in Four Countries. *American Journal of Sociology* 114(4): 977–1036.

Klatzky SR and Hodge RW (1971) A Canonical Correlation Analysis of Occupational Mobility. *Journal of the American Statistical Association* 66(333): 16–22.

Knoke D and Yang S (2008) *Social Network Analysis*. Thousand Oaks: SAGE Publications, Inc.

Levine JH (1972) A Two-Parameter Model of Interaction in Father-Son Status Mobility. *Behavioral Science* 17(5): 455–465.

MacDonald K (1972) MDSCAL and Distances Between Socio-Economic Groups. In: Hope K (ed.), *The Analysis of social mobility; methods and approaches*, Oxford studies in social mobility. Working papers, Oxford: Clarendon Press, pp. 211–234.

Marx K and Engels F (2008) *The communist manifesto*. London: Pluto Press.

Pakulski J (2005) Foundations of a post-class analysis. In: Wright EO (ed.), *Approaches to class analysis*, Cambridge, UK ; New York: Cambridge University Press, pp. 152–179.

Prandy K and Lambert P (2003) Marriage, Social Distance and the Social Space: An Alternative Derivation and Validation of the Cambridge Scale. *Sociology* 37(3): 397–411.

Savage M, Devine F, Cunningham N, et al. (2013) A New Model of Social Class? Findings from the BBC's Great British Class Survey Experiment. *Sociology* 47(2): 219–250.

Savage M, Devine F, Cunningham N, et al. (2015) On Social Class, Anno 2014. *Sociology* 49(6): 1011–1030.

Scott J (2000) *Social Network Analysis. A Handbook*. London: SAGE.

Stewart A, Prandy K and Blackburn RM (1980) *Social stratification and occupations*. New York: Holmes & Meier.

Treiman DJ (1977) *Occupational prestige in comparative perspective*. Quantitative studies in social relations, New York: Academic Press.

Toubøl J, Larsen AG and Jensen CS (2013) *A network analytical approach to the study of labour market mobility*. Paper presented at XXXIII Sunbelt Social Networks Conference of the International Network for Social Network Analysis (INSNA), Hamburg, Germany.

Toubøl J and Jensen CS (2014) Why do people join trade unions? The impact of workplace union density on union recruitment. *Transfer: European Review of Labour and Research* 20(1): 135-154.

Wacquant L (2013) Symbolic Power and Group-making: On Pierre Bourdieu's Reframing of Class. *Journal of Classical Sociology* 13(2): 274–291.

Walby S, Armstrong J and Strid S (2012) Intersectionality: Multiple Inequalities in Social Theory. *Sociology* 46(2): 224–240.

Weber M (1978) *Economy and society: an outline of interpretive sociology*. Berkeley: University of California Press.

Weeden KA and Grusky DB (2005) The Case for a New Class Map. *American Journal of Sociology* 111(1): 141–212.

Weeden KA and Grusky DB (2012) The Three Worlds of Inequality. *American Journal of Sociology* 117(6): 1723–1785.

Wright EO (1989) A General Framework for the Analysis of Class Structure. In: Wright EO (ed.), *The Debate on Classes*, London ; New York: Verso.

Wright EO (2005) Foundations of a neo-Marxist class analysis. In: Wright EO (ed.), *Approaches to class analysis*, Cambridge, UK ; New York: Cambridge University Press, pp. 4–30.

Author biographies

Jonas Toubøl is a PhD Fellow at Department of Sociology, University of Copenhagen. His research interests are in political sociology and social stratification, especially social movements, class and labour market segmentation.

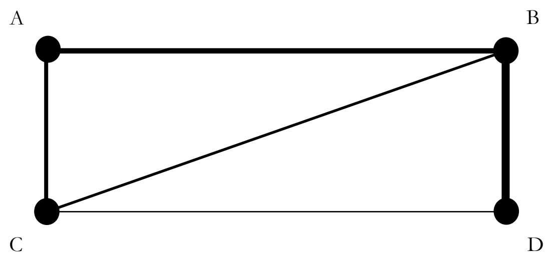
Anton Grau Larsen is Assistant Professor at Copenhagen Business School, Department of Business and Politics. His research interest are in elites, power, field analysis, class, and the methodologies of social network analysis and correspondence analysis.

Appendix A. Explaining and exemplifying The Mobility Network Clustering Algorithm

The arguments and figures in this appendix is a modified version of sections in Toubøl et al. (2013).

The MONECA algorithm is designed to identify discrete clusters of interconnected nodes in dense networks, as the one presented in figure A1. The figure depicts a graph of a rather dense weighted network as well as its adjacency table. The maximal cliques of the graph are |ABC| and |BCD|, which overlap each other in the case of 2 out of 3 nodes. In such a case, seeking to identify discrete clusters by a simple analysis of the cliques is a futile endeavour. MONECA offers a way to make sense of the discrete clustering structure of the network. In the following, we start by explaining the logic of MONECA followed by an example.

Figure A1: Exemplifying the MONECA algorithm



	A	B	C	D
A				
B	4			
C	3	2		
D	0	6	1	

Note: This figure has prior been presented at the XXXIII Sunbelt Social Networks Conference (Toubøl et al., 2013).

The logic of the algorithm is closely associated with the concept of the clique. The task of the algorithm is to determine to which clique to allocate the nodes that are in the overlapping areas of two or more cliques, as is the case of node B and C in figure A1. The network of figure A1 is weighted. The weights express the intensity of the relations of the nodes. This information enables the algorithm to decide which clique nodes B and C should belong to. Thus, in order to be able to produce a solution with MONECA, a network has to be weighted.

The algorithm is agglomerative, starting from the most disaggregated level, considering the connections of the single pairs of nodes. First step is to pair together the two most intensely connected nodes which then form a cluster. Subsequently, the connection between these two nodes is not considered. In step two MONECA proceeds to pair together the two nodes, which, among those remaining, are most intensely connected, in the same manner as in step one. And so it continues until all connections have been considered.

If one or both of the two most intensely connected nodes are already members of different clusters, they can only be joined together if all the nodes in the respective clusters are joined together, forming a new big cluster. However, a set of nodes can only be considered a cluster if they also form a clique. Therefore, in the case of pairing together two clusters, or adding a single node to an existing cluster, this is only possible if all the nodes under consideration form a clique. This criterion provides the stop rule for when no more single or sets of nodes should be paired together to form new clusters. Otherwise, the cluster solution would simply be the components of the network.

An example of how the algorithm works should clarify the procedure: In the case of figure A1, B and D are the most intensely connected nodes which can be seen from the width of the ties representing the intensity of the relation. Then, B and D are paired together, to form a preliminary cluster, |BD|. The second most intense connection is that of A and B. However, |BD| are already a cluster so MONECA asks whether A can be paired with both B *and* D, forming the cluster of |ABD|. In order to settle this question, MONECA must determine whether |ABD| constitutes a clique. In this case |ABD| is not a clique, because nodes A and D are not connected. Thus, nodes A and B cannot be paired. MONECA then goes on to consider the third strongest connection, which is A and C. Neither A or C are members of a preliminary cluster, and they can therefore be paired without further ado. We now have two clusters: |BD| and |AC|. The fourth strongest connection is BC. However, B has already been paired with D, and C has been paired with A. MONECA asks whether |ABCD| constitutes a clique. The answer is *no*, because A and D are not connected. Hence, |AC| and |BD| cannot be paired. The same is the case with regard to the fifth connection, |CD|. As result, the cluster solution produced by the algorithm is |AC| and |BD|, and none of the maximal cliques, |ABC|, |BCD|.

Appendix B. Cluster structure

		Level 1				Level 2				Level 3				Level 4				Level 5			
Occupation	*Final Cluster*	*Cluster*	*Within mob.*	*Density*	*Size*	*Cluster*	*Within mob.*	*Density*	*Size*	*Cluster*	*Within mob.*	*Density*	*Size*	*Cluster*	*Within mob.*	*Density*	*Size*	*Cluster*	*Within mob.*	*Density*	*Size*
832 Motor vehicle drivers	5.10	97	0.728	-	0.016	18	0.707	1.000	0.034	1	0.714	0.750	0.039	1	0.715	0.567	0.039	1	0.812	0.476	0.066
916 Garbage collectors & rel. lab.	5.10	105	0.491	-	0.001	18	0.707	1.000	0.034	1	0.714	0.750	0.039	1	0.715	0.567	0.039	1	0.812	0.476	0.066
931 Mining & construction lab.	5.10	107	0.647	-	0.017	18	0.707	1.000	0.034	1	0.714	0.750	0.039	1	0.715	0.567	0.039	1	0.812	0.476	0.066
811 Mining & mineral-proces.	5.10	80	0.545	-	0.000	24	0.539	1.000	0.005	1	0.714	0.750	0.039	1	0.715	0.567	0.039	1	0.812	0.476	0.066
833 agri. & other mobile plant op.	5.10	98	0.535	-	0.004	24	0.539	1.000	0.005	1	0.714	0.750	0.039	1	0.715	0.567	0.039	1	0.812	0.476	0.066
711 Miners, & stonecutters	5.10	64	0.583	-	0.000	-	-	-	-	-	-	-	-	1	0.715	0.567	0.039	1	0.812	0.476	0.066
712 Building frame & rel. trades wo.	5.10	65	0.820	-	0.028	-	-	-	-	-	-	-	-	-	-	-	-	1	0.812	0.476	0.066
733 Handicraft wo. in wood, textile	5.20	74	0.511	-	0.000	2	0.544	1.000	0.016	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
742 Wood treaters & rel. trades wo.	5.20	77	0.525	-	0.002	2	0.544	1.000	0.016	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
824 Wood products mach. op.	5.20	90	0.521	-	0.002	2	0.544	1.000	0.016	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
932 Manufacturing lab.	5.20	108	0.500	-	0.011	2	0.544	1.000	0.016	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
741 Food proces. & rel. trades wo.	5.20	76	0.592	-	0.005	10	0.649	1.000	0.020	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
827 Food & rel. products mach. op.	5.20	93	0.570	-	0.015	10	0.649	1.000	0.020	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
817 Automated assembly-line	5.20	86	0.575	-	0.000	13	0.544	1.000	0.015	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
821 Metal & mineral prod.	5.20	87	0.515	-	0.006	13	0.544	1.000	0.015	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
828 Assemblers	5.20	94	0.525	-	0.010	13	0.544	1.000	0.015	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
815 Chemical process. plant op.	5.20	84	0.560	-	0.002	15	0.492	1.000	0.008	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
822 Chemical products mach. op.	5.20	88	0.491	-	0.001	15	0.492	1.000	0.008	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
829 Other mach. op. & assemblers	5.20	95	0.449	-	0.005	15	0.492	1.000	0.008	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
812 Metal-process. plant op.	5.20	81	0.436	-	0.001	23	0.503	1.000	0.004	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
823 Rubber & plastic prod.	5.20	89	0.517	-	0.003	23	0.503	1.000	0.004	2	0.668	0.698	0.063	2	0.669	0.648	0.064	2	0.673	0.603	0.066
813 Glass, ceramics & rel. plant op.	5.20	82	0.512	-	0.000	-	-	-	-	-	-	-	-	2	0.669	0.648	0.064	2	0.673	0.603	0.066
743 Textile, garment wo.	5.20	78	0.585	-	0.001	12	0.576	1.000	0.003	-	-	-	-	-	-	-	-	2	0.673	0.603	0.066
826 Textile, fur & leather prod.	5.20	92	0.542	-	0.002	12	0.576	1.000	0.003	-	-	-	-	-	-	-	-	2	0.673	0.603	0.066
314 Ship & aircraft controllers	3.30	28	0.918	-	0.003	4	0.911	1.000	0.003	3	0.901	0.667	0.004	-	-	-	-	-	-	-	-
834 Ships' deck crews & rel. wo.	3.30	99	0.771	-	0.000	4	0.911	1.000	0.003	3	0.901	0.667	0.004	-	-	-	-	-	-	-	-
615 Fishery, hunters & trappers	3.30	63	0.733	-	0.001	-	-	-	-	3	0.901	0.667	0.004	-	-	-	-	-	-	-	-

235 Other teach. prof.	3.40	17	0.640	-	0.007	26	0.641	1.000	0.007	4	0.705	0.667	0.018	-	-	-	-	-	-	-	-
348 Religious ass. prof.	3.40	44	0.608	-	0.000	26	0.641	1.000	0.007	4	0.705	0.667	0.018	-	-	-	-	-	-	-	-
232 Sec. education teach. prof.	3.40	14	0.719	-	0.011	-	-	-	-	4	0.705	0.667	0.018	-	-	-	-	-	-	-	-
244 Social sciences & rel. prof.	3.50	21	0.635	-	0.007	1	0.770	1.000	0.013	5	0.786	0.600	0.023	-	-	-	-	-	-	-	-
344 Customs, tax	3.50	40	0.672	-	0.004	1	0.770	1.000	0.013	5	0.786	0.600	0.023	-	-	-	-	-	-	-	-
346 Social work ass. prof.	3.50	42	0.687	-	0.002	1	0.770	1.000	0.013	5	0.786	0.600	0.023	-	-	-	-	-	-	-	-
111 Legislators	3.50	1	0.618	-	0.001	5	0.740	1.000	0.009	5	0.786	0.600	0.023	-	-	-	-	-	-	-	-
242 Legal prof.	3.50	19	0.747	-	0.006	5	0.740	1.000	0.009	5	0.786	0.600	0.023	-	-	-	-	-	-	-	-
247 Adm. of legislation	3.50	24	0.569	-	0.002	5	0.740	1.000	0.009	5	0.786	0.600	0.023	-	-	-	-	-	-	-	-
332 Pre-primary education teach.	3.60	34	0.703	-	0.027	16	0.810	1.000	0.044	6	0.818	0.833	0.046	-	-	-	-	-	-	-	-
333 Special education teach.	3.60	35	0.644	-	0.015	16	0.810	1.000	0.044	6	0.818	0.833	0.046	-	-	-	-	-	-	-	-
334 Other teach. ass. prof.	3.60	36	0.522	-	0.002	16	0.810	1.000	0.044	6	0.818	0.833	0.046	-	-	-	-	-	-	-	-
331 Primary education teach.	3.60	33	0.621	-	0.002	-	-	-	-	6	0.818	0.833	0.046	-	-	-	-	-	-	-	-
413 Material-recording & transport	3.70	47	0.473	-	0.009	11	0.692	1.000	0.092	7	0.750	0.811	0.147	-	-	-	-	-	-	-	-
414 Library, mail & rel. Clerks	3.70	48	0.496	-	0.009	11	0.692	1.000	0.092	7	0.750	0.811	0.147	-	-	-	-	-	-	-	-
419 Other office clerks	3.70	49	0.347	-	0.004	11	0.692	1.000	0.092	7	0.750	0.811	0.147	-	-	-	-	-	-	-	-
522 Shop salespersons	3.70	58	0.634	-	0.045	11	0.692	1.000	0.092	7	0.750	0.811	0.147	-	-	-	-	-	-	-	-
911 Street vendors & rel. wo.	3.70	100	0.409	-	0.002	11	0.692	1.000	0.092	7	0.750	0.811	0.147	-	-	-	-	-	-	-	-
915 Messengers, porters	3.70	104	0.488	-	0.004	11	0.692	1.000	0.092	7	0.750	0.811	0.147	-	-	-	-	-	-	-	-
933 Transport lab. & freight	3.70	109	0.533	-	0.019	11	0.692	1.000	0.092	7	0.750	0.811	0.147	-	-	-	-	-	-	-	-
512 Housekeeping & restaurant	3.70	53	0.643	-	0.016	17	0.644	1.000	0.016	7	0.750	0.811	0.147	-	-	-	-	-	-	-	-
521 Fashion & other models	3.70	57	0.608	-	0.000	17	0.644	1.000	0.016	7	0.750	0.811	0.147	-	-	-	-	-	-	-	-
913 Domestic & rel. Helpers	3.70	102	0.612	-	0.039	-	-	-	-	7	0.750	0.811	0.147	-	-	-	-	-	-	-	-
721 Metal moulders, welders	3.80	68	0.622	-	0.009	14	0.752	1.000	0.037	8	0.802	0.850	0.064	-	-	-	-	-	-	-	-
722 Blacksmiths, toolmakers	3.80	69	0.602	-	0.011	14	0.752	1.000	0.037	8	0.802	0.850	0.064	-	-	-	-	-	-	-	-
723 Mach. mechanics & fitters	3.80	70	0.682	-	0.017	14	0.752	1.000	0.037	8	0.802	0.850	0.064	-	-	-	-	-	-	-	-
713 Building finishers	3.80	66	0.762	-	0.020	19	0.761	1.000	0.027	8	0.802	0.850	0.064	-	-	-	-	-	-	-	-
724 Electrical mechanics & fitters	3.80	71	0.513	-	0.007	19	0.761	1.000	0.027	8	0.802	0.850	0.064	-	-	-	-	-	-	-	-
422 Client information clerks	3.90	51	0.603	-	0.008	29	0.649	1.000	0.010	9	0.701	0.667	0.057	-	-	-	-	-	-	-	-
511 Travel attendants & rel. wo.	3.90	52	0.734	-	0.003	29	0.649	1.000	0.010	9	0.701	0.667	0.057	-	-	-	-	-	-	-	-
411 Secretaries	3.90	45	0.683	-	0.046	-	-	-	-	9	0.701	0.667	0.057	-	-	-	-	-	-	-	-
341 Finance & sales ass. prof.	3.10	37	0.683	-	0.039	28	0.702	1.000	0.054	10	0.713	0.667	0.059	-	-	-	-	-	-	-	-
412 Numerical clerks	3.10	46	0.587	-	0.006	28	0.702	1.000	0.054	10	0.713	0.667	0.059	-	-	-	-	-	-	-	-

421 Cashiers, tellers & rel. clerks	3.10	50	0.577	-	0.008	28	0.702	1.000	0.054	10	0.713	0.667	0.059	-	-	-	-	-	-	-	-
342 Business serv. Agents	3.10	38	0.675	-	0.005	-	-	-	-	10	0.713	0.667	0.059	-	-	-	-	-	-	-	-
311 Physical & engineering tech.	3.11	25	0.701	-	0.023	31	0.732	1.000	0.031	11	0.790	0.833	0.051	-	-	-	-	-	-	-	-
315 Safety & quality inspectors	3.11	29	0.604	-	0.002	31	0.732	1.000	0.031	11	0.790	0.833	0.051	-	-	-	-	-	-	-	-
321 Life science tech.	3.11	30	0.738	-	0.006	31	0.732	1.000	0.031	11	0.790	0.833	0.051	-	-	-	-	-	-	-	-
214 Architects & engineers	3.11	9	0.766	-	0.020	-	-	-	-	11	0.790	0.833	0.051	-	-	-	-	-	-	-	-
734 Printing & rel. trades wo.	3.12	75	0.621	-	0.003	3	0.697	1.000	0.006	12	0.680	0.667	0.006	-	-	-	-	-	-	-	-
825 Printing, binding & paper	3.12	91	0.594	-	0.003	3	0.697	1.000	0.006	12	0.680	0.667	0.006	-	-	-	-	-	-	-	-
814 Wood process. & papermaking	3.12	83	0.448	-	0.000	-	-	-	-	12	0.680	0.667	0.006	-	-	-	-	-	-	-	-
611 Market gardeners & crop	3.13	59	0.682	-	0.004	8	0.664	1.000	0.007	13	0.730	0.600	0.021	-	-	-	-	-	-	-	-
614 Forestry & rel. wo.	3.13	62	0.594	-	0.001	8	0.664	1.000	0.007	13	0.730	0.600	0.021	-	-	-	-	-	-	-	-
921 Agriculture, fishery & rel. lab.	3.13	106	0.528	-	0.002	8	0.664	1.000	0.007	13	0.730	0.600	0.021	-	-	-	-	-	-	-	-
612 Market-oriented animal prod.	3.13	60	0.625	-	0.005	32	0.752	1.000	0.014	13	0.730	0.600	0.021	-	-	-	-	-	-	-	-
613 Market-oriented crop & animal	3.13	61	0.738	-	0.010	32	0.752	1.000	0.014	13	0.730	0.600	0.021	-	-	-	-	-	-	-	-
212 Math & stat. prof.	3.14	7	0.529	-	0.000	27	0.706	1.000	0.050	14	0.723	0.650	0.066	-	-	-	-	-	-	-	-
241 Business prof.	3.14	18	0.637	-	0.018	27	0.706	1.000	0.050	14	0.723	0.650	0.066	-	-	-	-	-	-	-	-
343 Administrative ass. prof.	3.14	39	0.665	-	0.031	27	0.706	1.000	0.050	14	0.723	0.650	0.066	-	-	-	-	-	-	-	-
121 Directors & chief executives	3.14	2	0.679	-	0.005	30	0.648	1.000	0.016	14	0.723	0.650	0.066	-	-	-	-	-	-	-	-
123 Other departmental managers	3.14	4	0.614	-	0.011	30	0.648	1.000	0.016	14	0.723	0.650	0.066	-	-	-	-	-	-	-	-
245 Writers & performing artists	2.60	22	0.701	-	0.006	6	0.727	1.000	0.012	-	-	-	-	-	-	-	-	-	-	-	-
313 Optical & electronic equipment	2.60	27	0.686	-	0.003	6	0.727	1.000	0.012	-	-	-	-	-	-	-	-	-	-	-	-
347 Artistic, entertainment	2.60	43	0.618	-	0.003	6	0.727	1.000	0.012	-	-	-	-	-	-	-	-	-	-	-	-
213 Computing prof.	2.70	8	0.666	-	0.011	7	0.770	1.000	0.021	-	-	-	-	-	-	-	-	-	-	-	-
312 Computer ass. prof.	2.70	26	0.627	-	0.009	7	0.770	1.000	0.021	-	-	-	-	-	-	-	-	-	-	-	-
211 Nat. Science prof.	2.90	6	0.524	-	0.001	9	0.848	1.000	0.023	-	-	-	-	-	-	-	-	-	-	-	-
221 Life science prof.	2.90	10	0.562	-	0.002	9	0.848	1.000	0.023	-	-	-	-	-	-	-	-	-	-	-	-
222 Health prof. (except nursing)	2.90	11	0.922	-	0.013	9	0.848	1.000	0.023	-	-	-	-	-	-	-	-	-	-	-	-
231 Higher education teach. Prof.	2.90	13	0.613	-	0.007	9	0.848	1.000	0.023	-	-	-	-	-	-	-	-	-	-	-	-
223 Nursing & midwifery	2.20	12	0.725	-	0.005	20	0.860	1.000	0.029	-	-	-	-	-	-	-	-	-	-	-	-
323 Nursing & midwifery ass. prof.	2.20	32	0.822	-	0.024	20	0.860	1.000	0.029	-	-	-	-	-	-	-	-	-	-	-	-
233 Primary education teach. prof.	2.21	15	0.741	-	0.038	21	0.750	1.000	0.039	-	-	-	-	-	-	-	-	-	-	-	-
234 Special education teach. Prof.	2.21	16	0.468	-	0.001	21	0.750	1.000	0.039	-	-	-	-	-	-	-	-	-	-	-	-

122 Prod. & operations managers	2.22	3	0.567	-	0.012	22	0.617	1.000	0.024	-	-	-	-	-	-	-	-	-	-	-	-
131 General managers	2.22	5	0.427	-	0.012	22	0.617	1.000	0.024	-	-	-	-	-	-	-	-	-	-	-	-
816 Power prod.	2.25	85	0.679	-	0.001	25	0.652	1.000	0.011	-	-	-	-	-	-	-	-	-	-	-	-
914 Building caretakers	2.25	103	0.645	-	0.010	25	0.652	1.000	0.011	-	-	-	-	-	-	-	-	-	-	-	-
243 Librarians & rel. prof.	1.20	20	0.709	-	0.002	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
246 Religious prof.	1.23	23	0.935	-	0.001	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
322 Modern health ass. prof.	1.31	31	0.820	-	0.015	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
345 Police inspectors & detectives	1.41	41	0.893	-	0.001	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
513 Personal care & rel. wo.	1.54	54	0.740	-	0.102	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
514 Other personal service wo.	1.55	55	0.771	-	0.006	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
516 Protective serv. wo.	1.56	56	0.632	-	0.009	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
714 Painters	1.67	67	0.829	-	0.008	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
731 Precision wo. in metal	1.72	72	0.625	-	0.001	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
732 Potters, glass-makers	1.73	73	0.657	-	0.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
744 Leather & shoemaking	1.79	79	0.653	-	0.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
831 Locomotive engine-drivers	1.96	96	0.694	-	0.001	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
912 Shoe cleaning	1.10	101	0.500	-	0.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Appendix C: Reliability tests of MONECA

We perform three tests of the reliability of the methodological approach we have presented in this paper. These tests are concerned with the RR cut point of the minimum cell-frequency threshold, and sensitivity to changes in the data-input.

First, we test for the robustness with regard to the chosen cut point of an RR=1. This cut point is not arbitrary, as the assumption of randomness reflects the ideal of perfect mobility. Nonetheless, it is problematic for the reliability of the instrument if the cluster solution is sensitive to small variations in the cut point. On the other hand, we would expect large variation to significantly impact the solution.

Figure C1. Relative Risk cut point test results

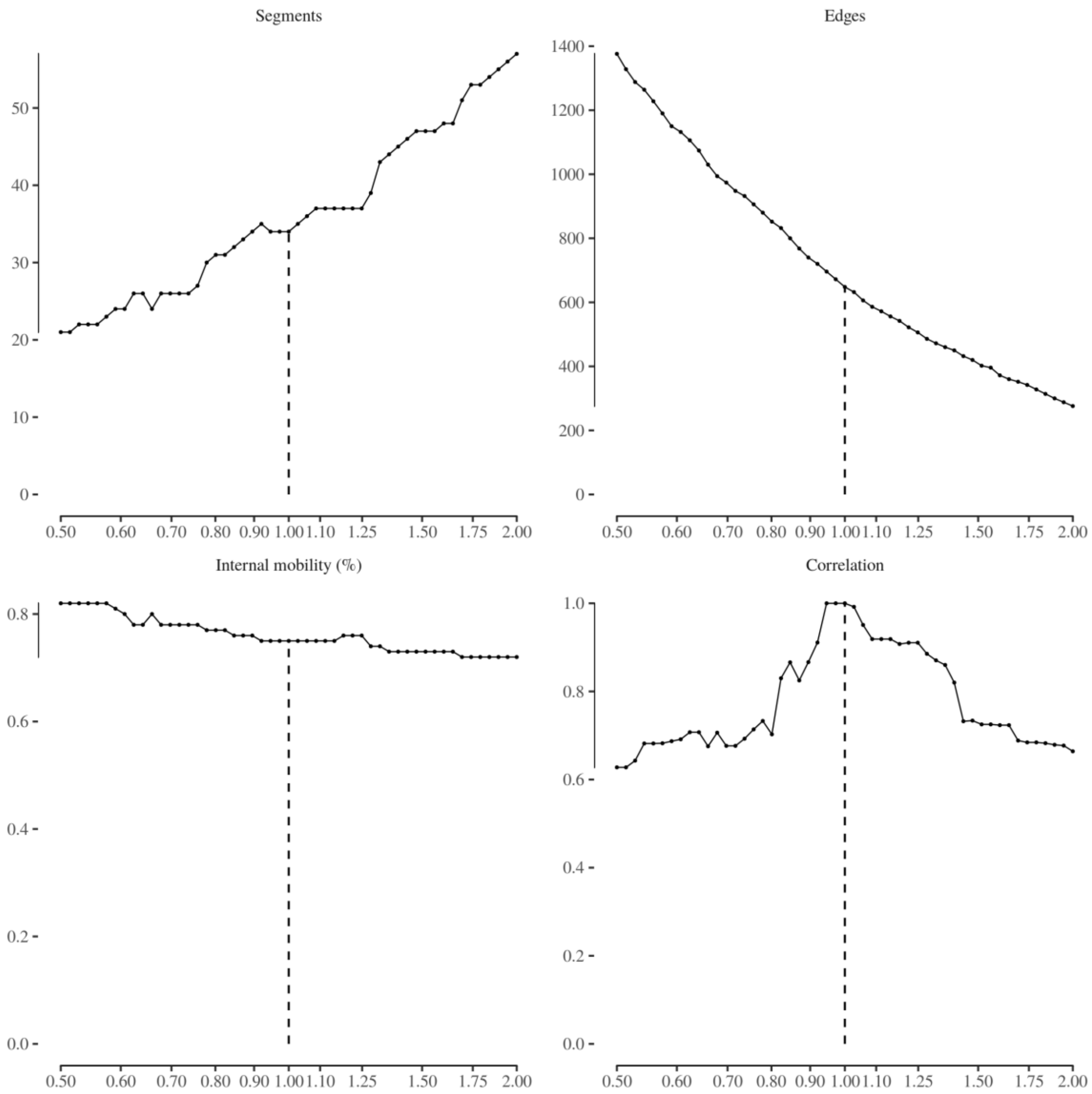
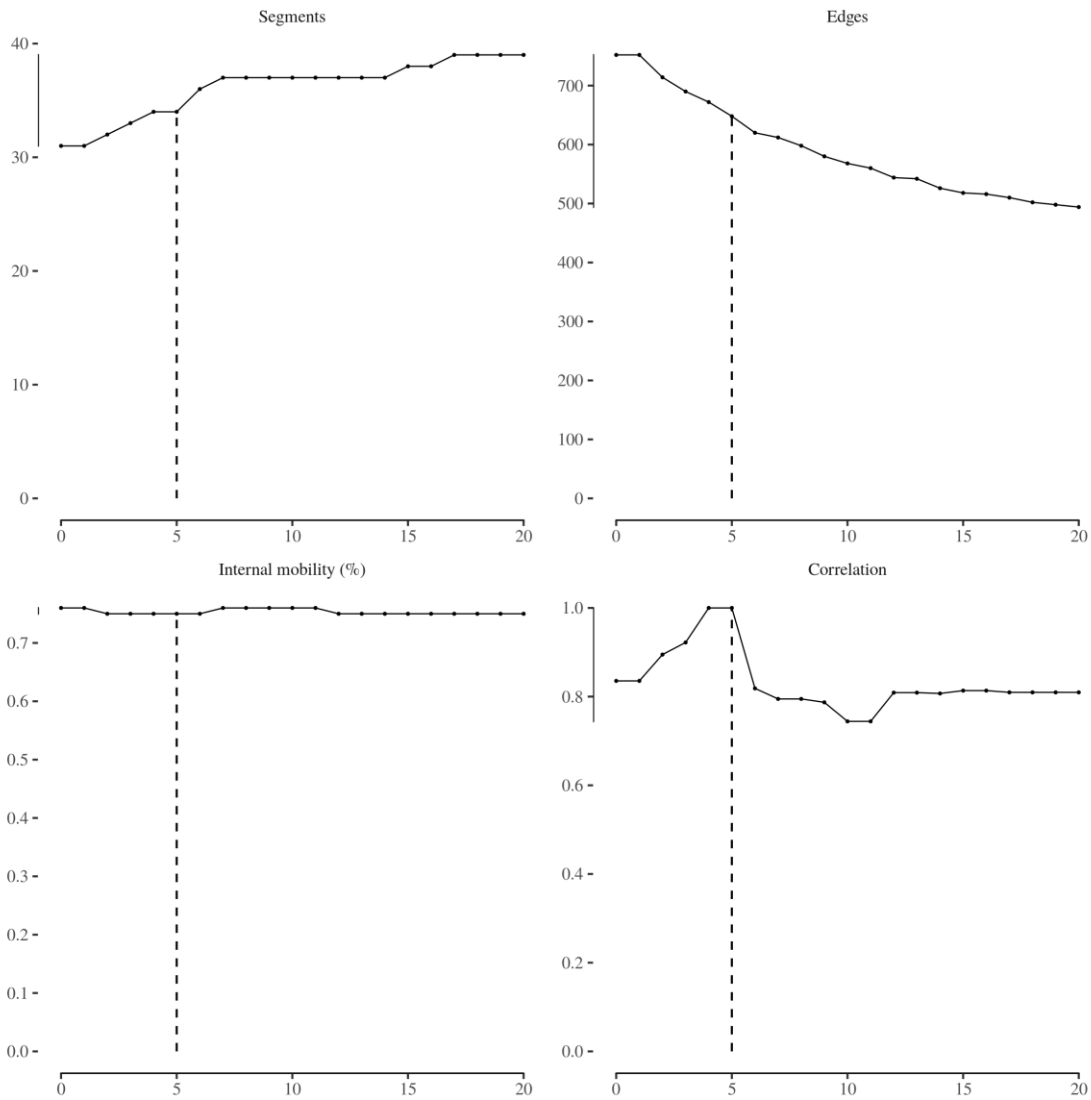


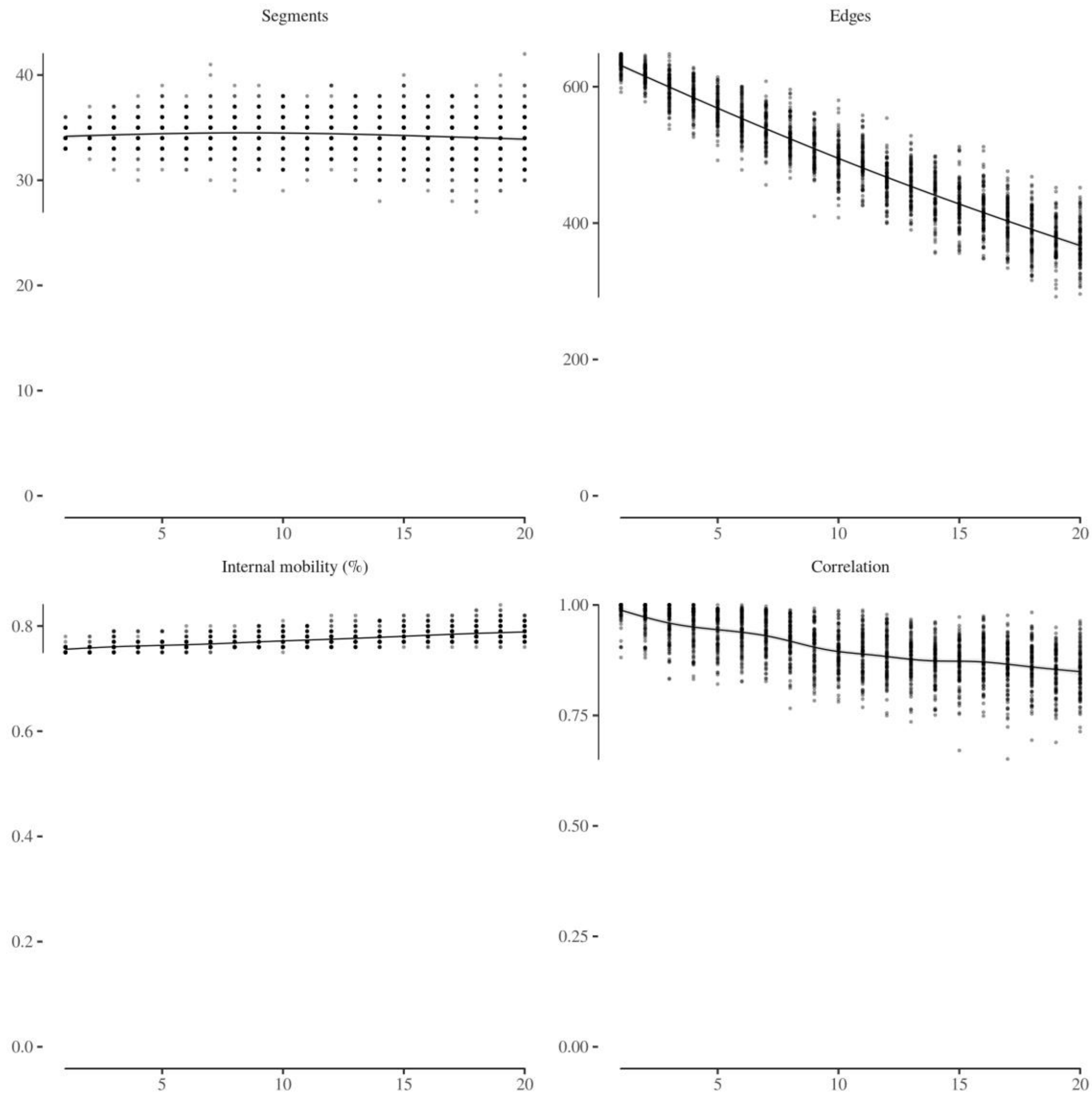
Figure C1 presents the results of the test of the robustness when varying the RR cut point. The four graphs show, respectively, 1) the change in number of edges, 2) the change in number of clusters of the final solution, 3) the amount of total mobility within the clusters of the final solution (i.e. mobility explained), 4) and the correlation of the cluster solution with the solution of an RR=1. The drop line indicates the solution when the RR cut point=1. As expected, the number of edges are affected dramatically by variation in the cut point. However, the number of clusters is quite stable around the original solution of 34 clusters, spanning from 30 to 37 within the 0.75≤RR≤1.2 interval. There is only little change with regard to the amount of mobility within the clusters. Finally, the correlation coefficient never goes below 0.6 within the entire test-interval of 0.5≤RR≤2.0. Within 0.825≤RR≤1.3 it never drops under 0.8, meaning 80% of the occupations are connected to, and separated from, the same clusters by virtue of the cluster boundaries of the final solution, when comparing with the solution of an RR=1. These results suggest that the algorithm is quite robust, despite the dramatic effect on the number of edges when varying the RR.

Figure C2. Minimum cell frequency threshold test results



Second, in the same manner we test the threshold for including ties of a minimum cell frequency of five. This threshold is more critical as there are no substantial theoretical considerations underpinning it, in contrast to what was the case with regard to the RR cut point. Thus, the choice of threshold rests on the researcher's inspection of what is necessary in order to exclude ties based on a relatively small number of observations that seem unreasonable according to substantial knowledge of the mobility structure under scrutiny. That is, we need to exclude ties based on a small number of observations, which we suspect to be the result of errors in measurement. This entails an aspect of subjectivity, and it would certainly be preferable if the derived cluster solutions are rather similar irrespective of this threshold. When inspecting figure C2, which is similar to figure C1 except for the X-axis representing variation in the minimum cell frequency threshold, we find only relatively small variation in the dependent measures. Thus, despite the subjective aspect in the choice of cut point, the algorithm produces very similar solutions, as is evident from the fact that the correlation with a solution of a threshold of 5 remains around 0.8 in the entire interval from 0 to 20.

Figure C3. Exclusion of nodes test results



Thirdly, we test the robustness of the MONECA algorithm by comparing the final solution with cluster solutions of networks in which we have deleted randomly chosen occupations. Figure C3 summarises the test results. This was done in order to assess the algorithm's ability to consistently reproduce the results when changing the underlying data structure. Even though it would be preferable to test by comparing the results of different datasets of labour market mobility, or different occupational classifications of the same data, the tests mimics such scenarios. In fact, it could be argued that removing parts of the table/network is a more challenging test than changing the data source or the classification, since the algorithm is based on the assumption that the data perfectly represents reality.

Because the mobility table is transformed into a network, removing any occupations/nodes may, in principle, affect the relations of all the remaining occupations/nodes. Thus, the changes to the data structure when deleting network nodes may be quite substantial. We deleted 1 to 20 occupations, and iterated the process 100 times. In total, 2000 reduced networks and subsequent cluster solutions were produced and compared with the solution of the full network. The full network consists of 109 occupations, meaning that removing 20 occupations is equal to deleting 18.3% of the nodes in the network. As can be seen from the lower right graph in figure C3, the average correlation is never lower than 0.9 and the lowest single correlation out of 2000 is at 0.65. This is reassuring with regard to the reliability of the cluster solution of the full network. It also suggests that MONECA may be rather robust across different data inputs varying in source and occupational classification.

Appendix D. Occupations by income, share of women, education and union density in 2007

Occupation	*Final cluster*	*Disp. income (DKK)*	*Education (Years)*	*Share of women*	*Union density*
733 Handicraft wo. in wood, textile & rel. mat.	5.20	179,692	11.5	44.9 %	79.6 %
741 Food proces. & rel. trades wo.	5.20	168,692	11.8	34.4 %	65.1 %
742 Wood treaters & rel. trades wo.	5.20	171,683	11.5	16.5 %	80.6 %
743 Textile, garment & rel. trades wo.	5.20	165,304	11.5	62.8 %	72.2 %
812 Metal-proces. plant op.	5.20	188,245	10.9	13.1 %	85.9 %
813 Glass. ceramics & rel. plant op.	5.20	201,708	10.8	15.3 %	89.1 %
815 Chemical proces. plant op.	5.20	219,256	11.6	25.0 %	89.9 %
817 Automated assembly-line & industrial robot op.	5.20	196,744	11.2	23.5 %	80.0 %
821 Metal & mineral products mach. op.	5.20	188,259	11.1	19.2 %	87.7 %
822 Chemical products mach. op.	5.20	192,938	11.1	23.8 %	81.7 %
823 Rubber & plastic products mach. op.	5.20	187,101	11.1	30.6 %	82.8 %
824 Wood products mach. op.	5.20	175,999	11.1	19.9 %	84.1 %
826 Textile. fur & leather products mach. op.	5.20	179,634	10.7	57.2 %	85.6 %
827 Food & rel. products mach. op.	5.20	192,858	11.2	31.2 %	92.1 %
828 Assemblers	5.20	179,960	11.1	53.6 %	87.0 %
829 Other mach. op. & assemblers	5.20	179,360	11.1	44.2 %	81.3 %
932 Manufacturing lab.	5.20	162,276	10.8	37.6 %	69.2 %
711 Miners. shot-firers. stonecutters & carvers	5.10	198,132	11.6	2.7 %	67.4 %
712 Building frame & rel. trades wo.	5.10	190,370	12.2	1.0 %	76.3 %
811 Mining & mineral-proces. plant op.	5.10	284,483	12.6	2.9 %	79.1 %
832 Motor vehicle drivers	5.10	204,378	10.8	5.7 %	79.0 %
833 Agri. & other mobile plant op.	5.10	200,352	10.6	4.7 %	84.2 %
916 Garbage collectors & rel. lab.	5.10	193,969	10.8	4.8 %	73.3 %
931 Mining & construction lab.	5.10	191,344	10.6	5.3 %	80.7 %
411 Secretaries & keyboard-operating clerks	3.90	194,636	12.7	83.9 %	74.7 %
422 Client information clerks	3.90	182,026	12.4	83.6 %	67.2 %
511 Travel attendants & rel. wo.	3.90	199,486	12.1	61.3 %	62.5 %
713 Building finishers & rel. trades wo.	3.80	198,297	12.6	1.1 %	81.1 %
721 Metal moulders. welders & rel. trades wo.	3.80	200,566	12.2	2.7 %	88.1 %
722 Blacksmiths. toolmakers & rel. trades wo.	3.80	204,120	12.5	2.8 %	87.1 %
723 mach. mechanics & fitters	3.80	212,477	12.6	2.5 %	83.3 %
724 Electrical mechanics & fitters	3.80	200,638	12.4	7.3 %	82.6 %
413 Material-recording & transport clerks	3.70	186,732	11.9	45.5 %	67.7 %
414 Library. mail & rel. Clerks	3.70	155,827	11.5	37.0 %	68.3 %
419 Other office clerks	3.70	148,883	12.1	68.2 %	53.5 %
512 Housekeeping & restaurant serv. wo.	3.70	140,114	11.5	65.1 %	50.5 %
521 Fashion & other models	3.70	132,978	11.9	67.4 %	23.4 %
522 Shop salespersons & demonstrators	3.70	115,193	11.2	59.6 %	40.0 %
911 Street vendors & rel. wo.	3.70	123,657	12.1	50.3 %	36.6 %
913 Domestic & rel. Helpers. cleaners & launderers	3.70	140,881	10.6	74.4 %	59.4 %
915 Messengers. porters. doorkeepers & rel. wo.	3.70	103,991	10.6	35.8 %	29.1 %
933 Transport lab. & freight handlers	3.70	132,202	10.7	27.1 %	48.6 %
331 Primary education teach. ass. prof.	3.60	209,019	13.9	93.5 %	93.6 %
332 Pre-primary education teach. ass. prof.	3.60	190,902	14.4	86.0 %	92.0 %
333 Special education teach. ass. prof.	3.60	190,741	14.1	76.0 %	87.3 %
334 Other teach. ass. prof.	3.60	192,953	13.1	47.2 %	69.9 %
111 Legislators	3.50	413,478	14.6	15.2 %	87.9 %
242 Legal prof.	3.50	312,539	16.7	53.0 %	88.7 %
244 Social sciences & rel. prof.	3.50	254,424	16.4	63.8 %	87.2 %
247 Adm. of legislation in public sector	3.50	270,004	15.5	49.8 %	91.6 %
344 Customs. tax & rel. gov. ass. prof.	3.50	219,146	13.4	79.0 %	92.4 %
346 Social work ass. prof.	3.50	213,602	14.7	82.0 %	89.7 %
232 Sec. education teach. prof.	3.40	279,033	15.8	44.3 %	93.2 %
235 Other teach. prof.	3.40	228,657	14.8	55.4 %	76.4 %
348 Religious ass. prof.	3.40	189,930	13.2	57.7 %	63.0 %

314 Ship & aircraft controllers & tech.*	3.30	338,274	13.2	10.8 %	60.1 %
615 Fishery wo.. hunters & trappers	3.30	346,832	11.1	4.2 %	45.2 %
834 Ships' deck crews & rel. wo.	3.30	231,192	11.2	3.7 %	45.5 %
121 Directors & chief executives	3.14	463,826	14.2	23.4 %	61.7 %
123 Other departmental managers	3.14	317,094	14.1	31.9 %	71.4 %
212 Math & stat. Prof.	3.14	325,706	16.2	35.4 %	61.1 %
241 Business prof.	3.14	303,843	14.7	43.2 %	65.1 %
343 Administrative ass. prof.	3.14	238,826	13.1	72.0 %	76.1 %
611 Market gardeners & crop growers	3.13	184,559	12.0	28.4 %	74.8 %
612 Market-oriented animal prod. & rel. wo.	3.13	320,369	11.9	33.3 %	47.7 %
613 Market-oriented crop & animal prod.	3.13	293,979	12.2	12.3 %	35.8 %
614 Forestry & rel. wo.	3.13	226,335	11.4	8.2 %	75.2 %
921 agriculture. fishery & rel. lab.	3.13	163,450	10.8	32.9 %	63.6 %
734 Printing & rel. trades wo.	3.12	228,120	12.8	29.3 %	79.3 %
814 Wood proces. & papermaking plant op.	3.12	190,686	11.2	21.0 %	84.1 %
825 Printing. binding & paper products mach. op.	3.12	218,735	11.6	15.8 %	91.2 %
214 Architects & engineers	3.11	299,033	15.8	19.9 %	74.7 %
311 Physical & engineering science tech.	3.11	244,970	13.5	29.0 %	79.6 %
315 Safety & quality inspectors	3.11	243,581	13.3	32.7 %	82.6 %
321 Life science tech. & rel. ass. prof.	3.11	222,353	14.5	78.5 %	85.2 %
341 Finance & sales ass. prof.	3.10	272,895	13.3	40.5 %	70.4 %
342 Business serv. agents & trade brokers	3.10	230,026	12.9	45.4 %	55.7 %
412 Numerical clerks	3.10	221,665	12.9	76.2 %	70.8 %
421 Cashiers. tellers & rel. Clerks	3.10	196,852	12.7	78.1 %	82.0 %
211 Nat. Science prof.	2.90	304,594	17.0	33.8 %	66.3 %
221 Life science prof.	2.90	265,153	17.1	47.9 %	79.8 %
222 Health prof. (except nursing)	2.90	351,334	17.1	55.0 %	88.9 %
231 Higher education teach. prof.	2.90	289,422	17.5	38.1 %	76.0 %
213 Computing prof.	2.70	297,456	14.3	20.0 %	61.8 %
312 Computer ass. prof.	2.70	263,205	13.5	21.7 %	68.5 %
245 Writers & performing artists	2.60	245,041	14.7	43.7 %	77.0 %
313 Optical & electronic equipment op.	2.60	222,804	13.4	43.7 %	55.7 %
347 Artistic. entertainment & sports ass. prof.	2.60	194,807	13.1	52.5 %	51.8 %
816 Power prod. & rel. plant op.	2.25	277,178	12.3	3.2 %	90.6 %
914 Building caretakers. window & rel. Cleaners	2.25	182,557	11.6	10.8 %	74.4 %
122 Prod. & operations managers	2.22	587,889	14.2	31.1 %	78.7 %
131 General managers	2.22	490,870	13.5	45.8 %	71.2 %
233 Primary education teach. prof.	2.21	215,816	15.3	66.6 %	85.5 %
234 Special education teach. prof.	2.21	247,283	15.3	65.2 %	91.2 %
223 Nursing & midwifery	2.20	250,498	15.8	96.3 %	95.2 %
323 Nursing & midwifery ass. prof.	2.20	212,113	14.9	95.3 %	92.8 %
831 Locomotive engine-drivers & rel. wo.	1.96	231,657	12.7	4.5 %	97.9 %
744 Leather & shoemaking trades wo.	1.79	206,694	11.8	38.8 %	70.5 %
732 Potters. glass-makers & rel. trades wo.	1.73	170,028	11.3	63.6 %	77.7 %
731 Precision wo. in metal & rel. mat.	1.72	192,225	12.5	55.4 %	65.1 %
714 Painters. building structure cleaners & rel. trade wo.	1.67	181,745	11.8	26.9 %	77.7 %
516 Protective serv. wo.	1.56	222,799	12.7	13.6 %	95.2 %
514 Other personal service wo.	1.55	145,392	12.0	86.3 %	63.3 %
513 Personal care & rel. wo.	1.54	164,446	11.6	87.8 %	77.0 %
345 Police inspectors & detectives	1.41	274,551	13.6	2.9 %	98.6 %
322 Modern health ass. prof. (except nursing)	1.31	191,830	13.8	91.3 %	88.7 %
246 Religious prof.	1.23	224,456	16.3	44.7 %	82.3 %
243 Librarians & rel. information prof.	1.20	215,744	14.9	75.1 %	83.6 %
912 Shoe cleaning & other street serv.	1.10	178,708	11.6	4.3 %	78.7 %
Total		201,538	12.7	51.5 %	74.1 %

* The very low mean years of education is due to a data error. Aircraft pilots and air traffic controllers are for reasons unknown registered with unrealistically short educations in the registers of Statistics Denmark (Albæk and Thomsen, 2011: 28).

Appendix E. Income statistics of clusters in 2007 (DKK)

#	Disposable income				Standard disposable income*	
	Mean	*1st quartile*	*Median*	*3rd quartile*	*Mean*	*Median*
2.22**	573,365	207,441	264,893	366,065	390,762	284,020
3.3	324,543	225,476	292,407	389,893	335,398	314,777
2.9	319,998	242,145	295,656	372,783	314,088	296,391
2.7	283,378	227,417	273,123	327,252	300,512	286,266
1.41	274,551	243,353	265,612	298,121	259,158	256,705
3.11	263,466	209,732	248,553	298,085	279,297	266,050
3.14	260,838	204,497	241,252	290,369	313,082	278,342
3.5	259,903	201,841	240,659	290,447	294,666	262,807
3.4	257,314	210,576	249,732	291,374	244,103	239,077
3.1	256,265	194,563	235,923	286,159	309,019	271,578
1.96	231,657	209,050	229,532	253,992	228,248	227,223
3.13	231,378	144,298	179,945	217,862	239,014	193,588
2.6	229,191	165,202	225,530	277,358	272,227	255,234
1.23	224,456	198,763	239,656	270,633	212,795	214,957
1.56	222,799	190,695	217,900	248,648	232,091	222,538
3.12	219,583	182,666	215,754	252,547	230,062	224,544
2.20	218,123	186,717	218,703	250,653	234,075	232,139
2.21	216,611	189,453	222,823	251,154	215,941	215,426
1.20	215,744	177,938	221,482	257,222	215,972	216,176
1.79	206,694	167,564	199,582	233,461	233,030	208,246
3.8	203,340	169,660	205,739	239,306	223,085	218,027
5.1	195,615	160,613	194,396	229,571	243,174	225,305
3.9	192,325	159,505	191,941	223,928	210,620	205,266
1.72	192,225	146,318	192,367	226,098	223,058	213,516
3.6	191,953	166,885	193,560	219,040	212,912	214,687
1.31	191,830	157,614	190,105	222,799	196,569	196,303
2.25	189,132	154,347	184,879	218,472	242,115	221,612
1.67	181,745	143,332	181,622	213,853	196,210	186,829
5.2	181,415	154,273	182,791	212,997	205,376	199,260
1.101	178,708	161,172	182,779	213,839	202,614	198,277
1.73	170,028	139,511	183,812	207,774	199,004	195,182
1.54	164,446	131,857	166,579	199,279	123,409	186,378
1.55	145,392	95,178	148,233	184,203	180,403	176,198
3.7	134,150	76,353	138,176	182,299	191,052	194,655
Total	201,561	151,927	196,964	241,250	200,689	195,403

* Standardization consists in calculating the mean for all full time employed male between 35-45 years old. Mean age for all is 40 (s.d. 3.14) and lowest and highest mean among the clusters are 39.2 and 41.7.
** The discrepancies between the various income and wage measures in the *2.22 Manager* cluster is due to the fact that there exist a small group of top income earners in the cluster, and the fact that much of the income is not from wage labour but from other sources like returns on capital. In fact, the highest incomes in this groups are not registered as full-time employed indicating that their income only to a limited extend comes from wage labour.